

Result No.	Query	Score	Match	Length	DB	ID	Description	
1		118.5	8.2	233	1	TNFA_CANFA	P51742 canis fami	
2		110.5	7.6	235	1	TNFA_RAT	P16599 rattus norv	
3		109.5	7.5	233	1	TNFA_FELCA	P13101 felis silve	
4		109.5	7.5	233	1	TNFA_MOUSE	Q06804 mus musculu	
5		108	7.4	235	1	TNFA_BOVIN	Q06599 bos taurus	
6		107.5	7.4	235	1	TNFA_PERLE	P36939 peromyscus	
7		107	7.4	193	1	TNFA_CAPHI	P13296 capra hircu	
8		107	7.4	232	1	TNFA_PIG	P23563 sus scrofa	
9		102.5	7.1	234	1	TNFA_SHEEP	P23383 ovis aries	
10		101.5	7.0	229	1	TNFA_CEREL	P51743 cervus elap	
11		100.5	6.9	205	1	TNFB_HUMAN	P01374 homo sapien	
12		99	6.8	234	1	TNFA_CAVPO	P51435 cavia porce	
13		98.5	6.8	233	1	TNFA_HUMAN	P01375 homo sapien	
14		97.5	6.7	281	1	FASL_HUMAN	P48023 homo sapien	
15		97	6.7	281	1	TRAI_HUMAN	P50591 homo sapien	
16		95	6.5	233	1	TNFA_TRIVU	P79374 trichosurus	
17		93.5	6.4	651	1	E2BD_YEAST	P12754 saccharomyc	
18		93.5	6.4	235	1	TNFA_RABIT	P04924 oryctolagus	
19		90.5	6.2	204	1	TNFB_PIG	P26445 sus scrofa	
20		89.5	6.2	234	1	TNFA_HORSE	P29553 equus cabal	
21		89.5	6.2	233	1	TNFA_PAPSP	P33620 papio sp. (
22		87.5	6.0	993	1	TSH_DROME	P22265 drosophila	
23		85.5	5.9	233	1	TNFA_MACEFA	P79337 macaca fasc	
24		85.5	5.9	233	1	TNFA_MACMU	P48094 macaca mula	
25		85.5	5.9	204	1	TNFB_BOVIN	Q06600 bos taurus	
26		85.5	5.9	358	1	VE2_HPV51	P26547 human papil	
27		85.5	5.9	351	1	YIIG_ECOLI	P32151 escherichia	
28		84.5	5.8	291	1	TRAI_MOUSE	P50592 mus musculu	
29		84	5.8	197	1	TNFB_RABIT	P10154 oryctolagus	
30		84	5.8	295	1	YFLF_CAUCR	P33976 caulobacter	
31		82.5	5.7	373	1	DHA_MCTU	P30234 mycobacteri	
32		82.5	5.7	680	1	NASP_RABIT	P27123 oryctolagus	
33		82.5	5.7	478	1	OCY2_PIG	Q28013 sus scrofa	
34		82.5	5.7	593	1	STX2_CANFA	Q28288 canis fami	
35		81.5	5.6	279	1	FASL_MOUSE	P41047 mus musculu	
36		81	5.6	278	1	FASL_RAT	P36940 rattus norv	
37		81	5.6	570	1	HAQ_NITEU	Q50925 bacterioma	
38		81	5.6	769	1	VPI_BPpH6	P11126 nictemiopha	
39		80.5	5.5	2067	1	BIMB_EMENI	P33144 emerice	
40		79.5	5.5	1150	1	2ACA_HUMAN	Q06190 homo sapien	
41		79.5	5.5	889	1	COPP_YEAST	P41811 saccharomyc	
42		79.5	5.5	2032	1	CTOG_HUMAN	Q14008 homo sapien	
43		79	5.4	561	1	BCHD_RHOCA	P26175 rhodobacter	

QY 60 CLTVSFYQVAALQGD LASLRAELQGHAEKLPAGAGAPKAGLEEPAVTAGLKIFEPPA 119

Db 32 CLSFLSLLVAGATTLLFCLLHFGVIGPQRELP-----NGQLISPLA 74
 QY 120 PGEGNSNSRKNK---RAYGCPETVTDCLQLIADSETPIQKGSYTFVPWLLSFKRGS 176
 Db 75 QTVKSSRTPSDKPVAVVANPE-----AEGQ-----LQWL--SRAN 110
 QY 177 AL-----EKENILVKETGYFFIYGOVLYTDKTYAMGHILQKVKHVFQ---DELSLV 227
 Db 111 ALLANGVELTDNQLVPSDGLYLIYQVLFKGGCGSTHVLTLHTISRFAVSTQTKVNL 170
 QY 228 TLFR--CIONMPETLPNNSCYS-----AGIAKLEEGDELQLAIPRENAQISLDGDTFFGA 281
 Db 171 SAIKSPQRETPREGTEAKPWYEPILYGGVQLEKGDRLSAEINLPNVLDFAESGQVYFGI 230
 QY 282 LKL 284
 Db 231 IAL 233
 RESULT 2
 TNFA_RAT
 ID TNFA_RAT STANDARD; PRT; 235 AA.
 AC P16599;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=TESTIS;
 RX MEDLINE; 94040766.
 RA KWON J., CHUNG I.Y., BENVENISTE E.N.;
 RL GENE 132:227-236(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA SHIRAI T., SHIMIZU N., HORIGUCHI S., ITO H.;
 RL AGRIC. BIOL. CHEM. 53:1733-1736(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92329007.
 RA ESTLER H.C., GREWE M., GAUSSLING R., PAVLOVIC M., DECKER K.;
 RL BIOL. CHEM. HOPPE-SEYLER 373:271-281(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA DECKER K.F.;
 RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [5]
 RP SEQUENCE OF 1-231 FROM N.A.
 RC TISSUE=TAIL;
 RA KIRISITS M.J., VARDIMON D., KUNZ H.W., GILL T.J. III;
 RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -!- SUBUNIT: HOMOTRIMER.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 EXTRACELLULAR SOLUBLE FORM.
 CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 PROTEOLYTIC PROCESSING.
 CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 AND MALNUTRITION.
 CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL; X66539; G395370; -;
 DR EMBL; L00981; G205254; -;

DR EMBL; D00475; G220921; -;
 DR EMBL; AJ002278; E1154446; -;
 DR EMBL; L19123; G310232; -;
 DR PIR; JU0029; JU0029.
 DR PIR; S21674; S21674.
 DR PIR; JN0868; JN0868.
 DR HSSP; P01375; ITNF.
 DR PROSITE; PS00251; TNF_1; 1.
 DR PROSITE; PS0049; TNF_2; 1.
 KW CYTOKINE; CYTOKININ; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 79
 FT CHAIN 80 235 TUMOR NECROSIS FACTOR.
 FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 148 179 BY SIMILARITY.
 FT CARBOHYD 86 86 POTENTIAL.
 FT CONFLICT 39 39 L -> P (IN REF. 3 AND 4).
 FT CONFLICT 163 163 I -> T (IN REF. 3 AND 4).
 FT CONFLICT 202 202 F -> S (IN REF. 3 AND 4).
 SQ SEQUENCE 235 AA; 25806 MW; CIDF18B7 CRC32;
 Query Match 7.6%; Score 110.5; DB 1; Length 235;
 Best Local Similarity 22.2%; Pred. No. 0.017;
 Matches 54; Conservative 45; Mismatches 87; Indels 57; Gaps 11;
 QY 60 CLTVSVFYVAALQGLASLRAELQG-HHAELKLPAGAPKAGLEAPAVTAGLKTFEPP 118
 Db 32 CLSFLSLLVAGATTLLFCLLHFGVIGPQRELPNG-----LPLISSMAQTLLR----- 82
 QY 119 ARGENSSNSRKNRAVQPEVTVDCLQLIADSETPIQKGSYTFVPWLLSFKRGSAL 178
 Db 82 -----SSONSDDKPVAVVANHQAEEQLEWLSQRANALLANG-----M 120
 QY 179 EKENILVKETGYFFIYGOVLYTDK---TYAMGHILQKVKHVFGEDELSTLTLFR--C 232
 Db 121 DLKDNQVVPADGLYLIYQVLFKGGCGPDYLLHTVTSRFAIS-YQEKVSLLSAISKSP 179
 QY 233 IONMPETLP-----NNSCYSAGIAKLEEGDELQLAIPRENAQISLDG--DVT-----FFGA 281
 Db 180 PKDTPGAEKLPWYEPILYGGVQLEKGDLL-----SAEVLNPKYLDITESGQVYFGV 232
 QY 282 LKL 284
 Db 233 IAL 235
 RESULT 3
 TNFA_FELCA
 ID TNFA_FELCA STANDARD; PRT; 233 AA.
 AC P19101;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS FELIS SILVESTRIS CATUS (CAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; CARNIVORA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RX MEDLINE; 91016860.
 RA MCGRAW R.A., COFFEY B.W., OTTO C.M., DREWS R.T., RAWLINGS C.A.;
 RL NUCLEIC ACIDS RES. 18:5563-5563(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RA DANIEL S.L., BRENNER C.A., LEGENDE A.M., SOLOMAN A., ROUSE B.T.;
 RL SUBMITTED (XXX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION

RNA 7:193-201(1986).

Query Match	7.1%	Score 102.5;	DB 1;	Length 234;																																																				
Best Local Similarity	22.3%;	Pred. No. 0.083;																																																						
Matches	56;	Conservative 41;	Mismatches 87;	Indels 65;	Gaps																																																			
QY	58	SC-CLTVVSYQVAA	LOGDLSRAELQGHHA	EKLPA	GAPAGKAGLEEA	PAVTA	GLK	KIFE	116																																															
DB	29	SCWCLSF	ELLVAGATT	FCLLHFG	VIGP	QRE-----	EQSP	---AGP	SFNR 72																																															
QY	117	PPAPG	CNCSN	SNKR	RAV	QGP	PE	TV	TQD	CLQL	IA	SE	PT	IO	KGS	YTF	VP	WLL	SF	KRS	176																																			
DB	73	PLV	OTL	RSS	Q	SN	K	KE	V	-----	HV	VA	IS	AP	Q	U	R	W	G	D	S	Y	A	N	A	L	M	A	---	N 116																										
QY	177	ALE	EN	K	N	I	L	V	K	E	T	G	Y	F	I	I	Q	V	L	Y	-----	T	D	K	T	Y	A	M	G	H	L	I	Q	K	V	H	V	F	G	E	L	S	L	T	Y	L	F	R	231							
DB	117	G	V	E	L	K	N	Q	L	V	P	T	D	G	L	Y	I	Y	S	Q	V	L	F	R	G	H	G	C	P	S	T	P	L	T	H	T	S	I	R	A	V	S	---	Y	O	T	K	V	N	I	L	S	A	I	K	175
QY	232	--	C	I	Q	N	N	P	E	L	P	N	-----	N	S	C	S	A	G	I	A	K	L	E	G	E	D	L	O	L	A	I	P	R	E	N	A	Q	I	S	L	-----	D	G	D	275										
DB	176	S	P	C	H	R	---	E	T	L	E	G	A	E	A	K	P	W	E	D	I	Y	O	G	G	V	O	L	E	K	D	R	L	-----	S	A	E	I	N	L	P	E	Y	D	A	E	S	G	225							
QY	276	V	T	F	F	G	A	L	K	L	284																																													
DB	226	Q	V	T	F	G	I	A	L	234																																														

RESULT	10
TNFA_CEREL	
ID	TNFA_CEREL STANDARD; PRT; 229 AA.
AC	P51743;
DT	01-OCT-1996 (REL. 34, CREATED)
DD	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE	TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN) (FRAGMENT).
GN	TNFA.
OS	CERVUS ELAPHUS (RED DEER).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; ARTIODACTYLA.
RN	[1]
RP	SEQUENCE FROM N.A.
RL	LOCKHART E.A.;
RL	SUBMITTED (SEP-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC	-1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS.
CC	-1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC	-1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC	-1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION.
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR	EMBL; U14683; G540279; -
DR	PROSITE; PS00251; TNF_1; 1.
DR	PROSITE; PS50049; TNF_2; 1.
KW	CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT	NON_TER 1
FT	PROPEP <1 73
FT	CHAIN 74 229
FT	TRANSEM 31 51
FT	DISULFID 141 173
SQ	SEQUENCE 229 AA; 249987 MW; 4FD25D97 CRC32;

Query Match 7.0%; Score 101.5; DB 1; Length 229;
Best Local Similarity 23.9%; Pred. No. 0.099;
Matches 60; Conservative 34; Mismatches 88; Indels 69; Gaps 13;

```

58 SC-CLTVSVSYQVAALQGDLASRAELQGHAEKLPAGAGAPKAGLEEPAPVANTGKIFE 111
   || || || || || || || || || || || || || || || || || || || ||
24 SCILSLFSLLVAGATTFLCLLHFGVIGPQRE-----EQSP---TGLSINS 67
   || || || || || || || || || || || || || || || || || || || ||
117 PPAPGEGNSSONSRNKRAVGPEETVTQDCQLQIADSETPTIQKGSYTFVPWLLS---- 172
   || || || || || || || || || || || || || || || || || || || ||
68 PLVOTLRSSQASINKRPAHVHVANINAG--QLL-----WLDSCANAL 108
   || || || || || || || || || || || || || || || || || || || ||
172 FKRGSALEEKENKILVKETGYFFIYQVLYTDKT-----YAMGHLIQRKKVHVFGDELSL 226
   || || || || || || || || || || || || || || || || || || || ||
109 MANGVKLE--DNQLVPTDGLYILYSQVLFGRGSCPSTPLFLTHLTISRIAVS-YQIKVNI 165
   || || || || || || || || || || || || || || || || || || || ||
227 VTLEFR--CIONMPETLP-----NNSCSYAGIAKUEGDEQLQIAPRNAQISL-----D 273
   || || || || || || || || || || || || || || || || || || || ||
166 LSAIKSPCHRETPWEAKPWEPIYQGGVFOLEKGDRL-----SAEINLPDYLDYAE 218
   || || || || || || || || || || || || || || || || || || || ||
QY 274 GDVTFEGALKL 284
   || || || || || || || || || || || || || || || || || || || ||
Db 219 SGQVYFGIALL 229
   || || || || || || || || || || || || || || || || || || || ||

RESULT 11
TNFB,HUMAN
ID TNFB,HUMAN STANDARD; PRT; 205 AA.
AC P01374;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN LTA OR TNFB.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 87217060.
RA NEDOSPASOV S.A., SHAKHOV A.N., TURETSKAYA R.L., METT V.A.,
RA AZIZOV M.M., GEORGEV G.P., KOROKO V.G., DOBRYNIN V.N.,
RA FILIPPOV S.A., BYSTROV N.S., BOLDYREVA E.F., CHUVPILO S.A.,
RA CHUMAKOV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RA COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 86086150.
RA NEDWIN G.E., JARRETT-NEDWIN J., SMITH D.H., NAYLOR S.L.,
RA SAKAGUCHI A.Y., GOEDEL D.V., GRAY P.W.;
RA J. CELL. BIOCHEM. 29:171-181(1985).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE; 87057135.
RA KOBAYASHI Y., MIYAMOTO D., ASADA M., OGINATA M., OSAWA T.;
RA J. BIOCHEM. 100:727-733(1986).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE; 85086243.
RA GRAY P.W., AGGARWAL B.B., BENTON C.V., BRINGMAN T.S., HENZEL W.J.,
RA JARRETT J.A., LEUNG D.W., MOFFAT B., NG P., SVEDERSKY L.P.,
RA PALLADINO M.A., NEDWIN G.E.;
RA NATURE 312:721-724(1984).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE; 92339500.
RA MATSUYAMA N., OKAWA N., TSUKII Y., ENDO T., KAJI A.;
RA FEBS LETT. 302:141-144(1992).
[6]
RN SEQUENCE FROM N.A.
RX MEDLINE; 9327029.
RA IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERINA D., PRIMAS G.,
RA PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA COHEN D.;
RA NAT. GENET. 3:137-145(1993).
[7]
RN PARTIAL SEQUENCE. CARBOHYDRATE-BINDING SITES. AND VARIANT.

```

RX MEDLINE: 93083656.
 RA VOIGT C.G., MAURER-FOGY I., ADOLF G.R.;
 RL FEBS LETT. 314:85-88(1992).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE: 92129275.
 RA ECK M.J., ULTSCH M., RINDERKNECHT E., DE VOS A.M., SPRANG S.R.;
 RL J. BIOL. CHEM. 267:2119-2122(1992).
 RN [9]
 RP VARIANT ALLELE 8.1.
 RX MEDLINE: 91139175.
 RA ABRAHAM L.J., DU D.C., ZAHEDI K., DAWKINS R.L., WHITEHEAD A.S.;
 RL IMMUNOGENETICS 33:50-53(1991).
 CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
 CC IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
 CC -1- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
 CC OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: X01393; G34445; -.
 DR EMBL: X02911; G312411; -.
 DR EMBL: D00102; D1000508; -.
 DR EMBL: M16441; G339740; -.
 DR EMBL: D12614; G219912; -.
 DR EMBL: M55913; G339743; -.
 DR EMBL: Z15026; G37213; -.
 DR EMBL: A06316; G412161; -.
 DR PIR: A26085; QWHUX.
 DR HSSP: P01375; 1TNR.
 DR MIW: 153440; -.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50045; TNF_2; 1.
 KW CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL; POLYMORPHISM.
 FT SIGNAL 1 34
 FT CHAIN 35 205 LYMPHOTOXIN-ALPHA.
 FT CARBOHYD 41 41 PARTIAL.
 FT CARBOHYD 96 96
 FT VARIANT 60 60 T -> N.
 FT VARIANT 125 125 T -> P (IN ALLELE 8.1).
 FT CONFLICT 13 13 C -> R (IN REF. 6).
 SQ SEQUENCE 205 AA; 22297 MW; 42F296D8 CRC32;

Query Match 6.9%; Score 100.5; DB 1; Length 205;
 Best Local Similarity 23.6%; Pred. No. 0.1;
 Matches 51; Conservative 20; Mismatches 84; Indels 61; Gaps 9;

QY 91 LPAGAGAFKAGLEAPAVTA-----GLKIFEPAPGEGN--SSONSRRKRAVGQPEET 141
 DB 29 LPGAGLPGVGLTPSAQTARQHPKMHLSHTLKPAAHLIGDPKSNLLWRA--NTDRA 86
 QY 142 VTQDCLQIADSETPTIOKSYTFVFWLLSPKRGSALEKENKILVKETGYFFIYGQVLY 201
 DB 87 FLQDQFSL-----SNNLLVPTSGIYFYVSQVFF 115
 QY 202 TDKIYA----MGHLIQRKKVHVFGDELSLVTLFRCIQNMPTLPN-----NSCYSAGIA 251
 DB 116 SKAYSPRKATSPSYLAHEVOLFSQYFHPVLLSSQKN--VYFGLQEPWLHSMYHGAFF 173
 QY 252 KLEEGDELQQL--AIPRENAOISLDGDTFFFGALKL 284
 DB 174 QLTQDQLSTHTDGP-----HLVLSPTVFFGAFAL 205

RESULT 12
 TNFA_CAVPO STANDARD; PRT; 234 AA.
 AC P51435;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.

OS CAVIA PORCELLUS (GUINEA PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY; TISSUE=LUNG;
 RA YUAN H.T., KELLY F.J., BINGLE C.D.;
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DUNKIN-HARTLEY;
 RA WHITE A.M., YOSHIMURA T., SMITH A.W., WESTWICK J., WATSON M.L.;
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
 CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CC CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING (BY SIMILARITY).
 CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 CC AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: U39839; G1066112; -.
 DR EMBL: U77036; G1679724; -.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 79 BY SIMILARITY.
 FT CHAIN 80 234 TUMOR NECROSIS FACTOR.
 FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 147 178 BY SIMILARITY.
 SQ SEQUENCE 234 AA; 25793 MW; 00160783 CRC32;

Query Match 6.8%; Score 99; DB 1; Length 234;
 Best Local Similarity 22.3%; Pred. No. 0.17;

Matches 55; Conservative 41; Mismatches 85; Indels 66; Gaps 13;

QY 60 CLTVVSFYQVAALQGLASRLAEQLQGHAEKLPAGAKGAEAPAVTAGLKITEPPA 119
 DB 32 CLSLFELLVAGATTFLCLLHFGVIGPQREE-QFSSGPPFRPL-----AQTLTLR----- 81
 QY 120 PGENSSONSRRKRAVGQPEETVTQDCLQIADSETPTIOKSYTFVFWLLSPKRGSALE 179
 DB 81 ----SASQNDNDKPVHVAVANQAEEELQ-----WL--SKRANALL 115
 QY 180 E-----KENKILVKETGYFFIYGQVLYTDK---TY-AMGHLIQRKKVHVFGDELSLVTLF 230
 DB 116 ANGMGLSDNOLVPSDGLYLIYSQVLFKGCPCPSYLLLTHTVSRLAWS-YPEKVNLLSAI 174
 QY 231 R--CIONMPTPLNNSCYS---AGIAKLEEGDELQALAIPRENAOISL-----DGDVT 277
 DB 175 KSPCKETPEGAERKPWFYFIYLGGVFLQKGDRLL-----SAENVLPQYLOFADFSGQI 227
 QY 278 FFGALKL 284
 DB 228 YFGVIAL 234

RESULT 13
 TNFA_HUMAN STANDARD; PRT; 233 AA.
 ID TNFA_HUMAN
 AC P01375;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE: 87217060.
RA NEDOSPASOV S.A., SHAKHOV A.N., TURETSKAYA R.L., METT V.A.,
RA AZIZOV M.M., GEORGIEV G.P., KOROKO V.G., DOBRYNIN V.N.,
RA FILIPPOV S.A., BYSTROV N.S., BOLDYREVA E.F., CHUVPILO S.A.,
RA CHUMAKOV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RL COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE: 85086244.
RA PENNICA D., NEDWIN G.E., HAYFLICK J.S., SEEBURG P.H., DERYNCK R.,
RA PALLADINO M.A., KOHR W.J., AGGARWAL B.B., GOEDDEL D.V.;
RL NATURE 312:724-729(1984).
RN [3]
RX SEQUENCE FROM N.A.
RX MEDLINE: 85137898.
RA SHIRAI T., YAMAGUCHI H., ITO H., TODD C.W., WALLACE R.B.;
RL NATURE 313:803-806(1985).
RN [4]
RX SEQUENCE FROM N.A.
RX MEDLINE: 86016093.
RA NEDWIN G.E., NAYLOR S.L., SAKAGUCHI A.Y., SMITH D.H.,
RA JARRETT-NEDWIN J., PENNICA D., GOEDDEL D.V., GRAY P.W.;
RL NUCLEIC ACIDS RES. 13:6361-6373(1985).
RN [5]
RX SEQUENCE FROM N.A.
RX MEDLINE: 8512190.
RA WANG A.M., CREASEY A.A., LADNER M.B., LIN L.S., STRICKLER J.,
RA VAN ARSDELL J.N., YAMAMOTO R., MARK D.F.;
RL SCIENCE 228:149-154(1985).
RN [6]
RX SEQUENCE FROM N.A.
RX MEDLINE: 86030296.
RA MARMEOUT A., FRANSEN L., TAVERNIER J., DER HEYDEN J., TIZARD R.,
RA KAWASHIMA E., SHAW A., JOHNSON M.J., SEMON D., MUELLER R.,
RA RUYSSCHAERT M.R., VAN VLIET A., FIERIS W.;
RL EUR. J. BIOCHEM. 152:515-522(1985).
RN [7]
RX SEQUENCE FROM N.A.
RX MEDLINE: 93272029.
RA IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERINA D., PRIMAS G.,
RA PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA COHEN D.;
RL NAT. GENET. 3:137-145(1993).
RN [8]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE: 89159409.
RA JONES E.Y., STUART D.I., WALKER N.P.;
RL NATURE 338:225-228(1989).
RN [9]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE: 91193276.
RA JONES E.Y., STUART D.I., WALKER N.P.;
RL J. CELL SCI. SUPPL. 13:11-18(1990).
RN [10]
RX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE: 90008932.
RA ECK M.J., SPRANG S.R.;
RL J. BIOL. CHEM. 264:17595-17605(1989).
RN [11]
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MUTANT ARG-107.
RX MEDLINE: 98147459.
RA REED C., FU Z.Q., WU J., XUE Y.N., HARRISON R.W., CHEN M.J.,
RA WEBER I.T.;
RL PROTEIN ENG. 10:1101-1107(1997).
RN [12]
RX MUTAGENESIS.

RX MEDLINE: 91184128.
RA OSTADE X.V., TAVERNIER J., PRANGE T., FIERIS W.;
RL EMBO J. 10:827-836(1991).
RN [13]
RX MYRISTOYLATION.
RX MEDLINE: 93018820.
RA STEVENSON F.T., BURSTEN S.L., LOCKSLEY R.M., LOVETT D.H.;
RL J. EXP. MED. 176:1053-1062(1992).
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS. IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X02910; G37210; -.
DR EMBL: M16441; G339741; -.
DR EMBL: X01394; G37220; -.
DR EMBL: M10988; G339738; -.
DR EMBL: M26331; G339784; -.
DR EMBL: Z15026; G37212; -.
DR PIR: B23784; QWHUN.
DR PIR: A44189; A44189.
DR PIR: S36153; S36153.
DR PDB: 1TNF; 15-JAN-91.
DR PDB: 2TUN; 31-JAN-94.
DR PDB: 1A8M; 17-JUN-98.
DR MIM: 191160; -.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR;
FT MYRISTYLATION: 3D-STRUCTURE.
FT PROPEP 1 76
FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT LIPID 19 19 MYRISTATE.
FT LIPID 20 20 MYRISTATE.
FT DISULFID 145 177
FT MUTAGEN 105 105 L->S: LOW ACTIVITY.
FT MUTAGEN 108 108 R->W: BIOLOGICALLY INACTIVE.
FT MUTAGEN 112 112 L->F: BIOLOGICALLY INACTIVE.
FT MUTAGEN 160 160 A->V: BIOLOGICALLY INACTIVE.
FT MUTAGEN 162 162 S->F: BIOLOGICALLY INACTIVE.
FT MUTAGEN 167 167 V->A,D: BIOLOGICALLY INACTIVE.
FT MUTAGEN 222 222 E->K: BIOLOGICALLY INACTIVE.
FT CONFLICT 63 63 F -> S (IN REF. 5).
FT TURN 85 86
FT STRAND 89 94
FT TURN 99 100
FT STRAND 104 105
FT TURN 109 110
FT STRAND 112 113
FT TURN 115 116
FT STRAND 118 119
FT STRAND 124 125
FT STRAND 132 138
FT STRAND 140 142
FT STRAND 152 159
FT STRAND 166 174
FT STRAND 190 191
FT STRAND 194 200
FT STRAND 207 212
FT HELIX 215 217
FT STRAND 218 218

SUBUNIT (MAY 1997) TO EMBL/GENBANK/DDSB DATA BANKS.
 -!- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
 TRANSDUCE THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
 CYTOTOXIC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
 FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
 PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
 T CELLS, OR BOTH
 -!- SUBUNIT: HOMOTRIMER (PROBABLE).
 -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
 INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
 SURFACE.
 CC
 CC
 -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 EMBL; X89102; G887456; -.

RESULT 15	
TRAIL_HUMAN	
ID	TRAIL_HUMAN
AC	STANDARD;
DT	PRT; 281 AA.
DT	01-OCT-1996 (REL. 34, CREATED)
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE	TNF-RELATED APOPTOSIS INDUCING LIGAND (TRAIL PROTEIN) (APO-2 LIGAND).
OS	TRAIL.
GN	HOMO SAPIENS (HUMAN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; PRIMATES.

515

```
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: WIDESPREAD; MOST PREDOMINANT IN SPLEEN, LUNG
CC AND PROSTATE.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; U37518; G1149558; -
DR EMBL; U57059; G1336208; -
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS0049; TNF_2; 1.
KW CYTOKINE; TRANSMEMBRANE; SIGNAL-ANCHOR; APOPTOSIS.
FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 18 38 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 39 281 EXTRACELLULAR (POTENTIAL).
SQ SEQUENCE 281 AA; 32509 MW; 4D80D8B1 CRC32;

Query Match          6.7%; Score 97; DB 1; Length 281;
Best Local Similarity 20.6%; Pred. No. 0.31;
Matches 37; Conservative 39; Mismatches 52; Indels 52; Gaps 9;

QY 121 GEGN--SSQNSRNKRAVGPEETVTQDCQLIADSETPTIQKGYTFVPWLLSFKRGSAL 178
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 131 GRSNTLSSPNSKNEKALG-----RKINSWES---SRSGHSFL-----SNL 167

QY 179 EKENKILVKETGYFFIYQVLYTDKTYAMGHLIQKKVHVFGDELSTVLFRCIQNMPE 238
    : : : : | | : | | : | | : | | : | | : | | : | | : | |
Db 168 HLRNGELVIHERKGFYIYSQTVF-----RFQEEIKENTKNDKQMQVIYK-YTSYPD 218

QY 239 TL-----PNNSC-----YSAGIAKLEEGDELQLAIPRENAQISLDGDVTFFGA 281
    : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 PILLMKSARNCSWSDAEYGLYSIYOGGIFELKENDRIFFSVTNEHL-IDMDHEASFFGA 277
```

Search completed: May 19, 1999, 14:43:36
Job time: 43 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:20:26 ; Search time 33.52 Seconds
(without alignments)
318.502 Million cell updates/sec

Title: US-09-212-270-1
Perfect score: 1451
Sequence: 1 MDDSTEREQSLTSCLKRE.....ENAIQISLDGVDVTFGALKLL 285

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR_58:*

- 1: PIR1.*
- 2: PIR2.*
- 3: PIR3.*
- 4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110.5	7.6	235	2 JU0029	tumor necrosis fac
2	109.5	7.5	235	1 QWMSN	tumor necrosis fac
3	109.5	7.5	233	2 S11688	tumor necrosis fac
4	108	7.4	233	2 S24642	tumor necrosis fac
5	107.5	7.4	235	2 I54490	tumor necrosis fac
6	107	7.4	232	2 S21606	tumor necrosis fac
7	107	7.4	193	2 S06192	tumor necrosis fac
8	106	7.3	233	2 S13114	tumor necrosis fac
9	102.5	7.1	234	2 JH0529	tumor necrosis fac
10	100.5	6.9	205	1 QWHUX	lymphotoxin alpha
11	98.5	6.8	233	1 QWHUN	tumor necrosis fac
12	98	6.8	234	2 A25451	tumor necrosis fac
13	97.5	6.7	281	2 I38707	Fas ligand - human
14	93.5	6.4	651	1 RBYD2	translation regula
15	92.5	6.4	185	2 S52715	tumor necrosis fac
16	90.5	6.2	204	2 S17289	tumor necrosis fac
17	89.5	6.2	233	2 S22052	tumor necrosis fac
18	89.5	6.2	234	2 J01344	tumor necrosis fac
19	87.5	6.0	993	2 A38437	probable homeotic
20	85.5	5.9	358	1 W3WL51	E2 protein - human
21	85.5	5.9	204	2 S24641	lymphotoxin - bovi
22	85.5	5.9	351	2 S40840	hypothetical 39.3K
23	84	5.8	197	2 JH0309	tumor necrosis fac
24	84	5.8	295	2 B41320	hypothetical prote
25	82.5	5.7	371	1 A43830	alanine dehydrogen
26	82.5	5.7	478	1 I47154	transcription fact
27	82.5	5.7	665	2 S28182	lamin A - mouse
28	82.5	5.7	574	2 S04333	lamin C - mouse
29	82.5	5.7	760	2 S55520	Chitin synthetase
30	82.5	5.7	680	2 A43800	nuclear autoantige
31	82	5.7	1286	2 B71433	hypothetical prote
32	81.5	5.6	562	2 S75308	DNA ligase (BC 6.5
33	81.5	5.6	279	2 A53062	Fas ligand - mouse
34	81	5.6	769	1 P1BP66	P1 protein - phage
35	81	5.6	570	2 A36954	hydroxylamine oxid
36	81	5.6	613	2 S27770	hypothetical prote
37	81	5.6	398	2 S40236	fetal beta-MHC-bin
38	81	5.6	278	2 A49266	fasc ligand - rat
39	80.5	5.5	2067	2 A42854	probable spindle p

40 80 5.5 905 2 S55059 fertilin alpha-I -
41 80 5.5 886 2 I38182 transcription term
42 79.5 5.5 26926 1 I38344 titin, cardiac mus
43 79.5 5.5 889 2 B55123 coatomer complex b
44 79.5 5.5 1150 2 B47114 phosphoprotein pho
45 79.5 5.5 1972 2 S68176 TOG protein - huma

ALIGNMENTS

RESULT 1
JU0029
tumor necrosis factor alpha precursor - rat
N:Alternate names: cachectin; TNF alpha
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Mar-1998
C:Accession: JU0029; JN0868; S21674
R:Shirai, T.; Shimizu, N.; Horiguchi, S.; Ito, H.
Agric. Biol. Chem. 53, 1733-1736, 1989
A:Title: Cloning and expression in Escherichia coli of the gene for rat tumor necrosi
A:Reference number: JU0029
A:Accession: JU0029
A:Molecule type: DNA
A:Residues: 1-235 <SHI>
R:Kwon, J.; Chung, I.Y.; Benveniste, E.N.
Gene 132, 227-236, 1993
A:Title: Cloning and sequence analysis of the rat tumor necrosis factor-encoding gene
A:Reference number: JN0868
A:Accession: JN0868
A:Molecule type: DNA
A:Residues: 1-235 <KW>
A:Cross-references: GB:L00981; NID:g205253; PID:g205254
R:Estler, H.C.; Grewe, M.; Gaussling, R.; Pavlovic, M.; Decker, K.
Biol. Chem. Hoppe-Seyler 373, 271-281, 1992
A:Title: Rat tumor necrosis factor-alpha. Transcription in rat Kupffer cells and in v
A:Reference number: S21674; MUID:92329007
A:Accession: S21674
A:Molecule type: mRNA
A:Residues: 1-38, 'P', 40-162, 'T', 164-201, 'S', 203-235 <EST>
A:Cross-references: GB:X66539; GB:S40199; NID:g995369; PID:g995370
C:Comment: Tumor necrosis factor is secreted by macrophages in response to endotoxin
C:Genetics:
A:Gene: TNF-alpha
A:Introns: 62/3: 81/1; 97/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
F:80-235/Product: tumor necrosis factor #status predicted <MAT>
F:86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	7.6%	Score	110.5;	DB	2;	Length	235;
Best Local Similarity	22.2%	Pred. No.	0.019;				
Matches	54;	Conservative	45;	Mismatches	87;	Indels	57;
Gaps	11;						
Qy	60	CLTVVSEYQVAALQGLASRAELQG-HHAELKLPAGACAPKAGLEAPANTAGLAIFFPP	118				
Db	32	CLSLFSLLVAGATTFLCFLNFGVIGPNKEKFPNG-----LPLISSMAQTTLR-----	82				
Qy	119	APGEGNSSQNRKRAVQGPETVTQDCLOLIADSETPTIQKSYTFVPWLLSFKRGSA	178				
Db	82	-----SSQNSDKPVAHVANHQAEQLEWLSQANALLANG-----M	120				
Qy	179	EKENKILVKTGYFFIYGVLYTDK-----TYAMGHLIQRKKVHVFGDELSVTLFR--C	232				
Db	121	DLKDNQLVVPADGLYLIYSVLFKGGQCPDVLLTHTVTSRFAIS-YQEKVLSLSA	179				
Qy	233	IQNPETLP-----NNSCYSAGIAKLEEGDELQALPFRNAQISLDG--DVT-----	281				
Db	180	PROTFEGAEALPKWPMPYMLGGVFOLEKGDLL-----SAEVLNPKYLDIT	232				
Qy	282	LKL	284				

Db 176 PCQRETEGAERAKPWYPIYLGQVQLEKGRDLSTEI---NLPAYLDFAESGV-YFGII 231
Qy 283 KL 284
Db 232 AL 233

RESULT 4
S24642
tumor necrosis factor alpha - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Feb-1997
C:Accession: I46047; S24642
R:Clufts, I.; Cleuter, Y.; Kettmann, R.; Burny, A.; Droogmans, L.
Cytochrome 5, 336-341, 1993
A:Title: Cloning and characterization of the tandemly arranged bovine lymphotoxin and tu
A:Reference number: I46046; MUID:94083525
A:Accession: I46047
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-233 <CL>
A:Cross-references: EMBL:Z14137; NID:g796; PID:g798
C:Genetics:
A:Gene: TNF
A:Introns: 62/3; 78/1; 94/1
C:Superfamily: tumor necrosis factor
C:Keywords: transmembrane protein

Query Match 7.4%; Score 108; DB 2; Length 233;
Best Local Similarity 22.8%; Pred. No. 0.031;
Matches 56; Conservative 42; Mismatches 88; Indels 60; Gaps 12;
Qy 58 SC-CLTVVSFYQVAALQGLDASLRAELQGHAEKLPAGAGAPKAGLEAPAVTAGLKIFE 116
Db 29 SCLCLSLFLLVAGATTFLCLLHFGVIGPQREESPG-----PSINS----- 72
Qy 117 PPAPGEGNSONSNNKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPLLSFKRGS 176
Db 72 PLVQTLRSSQASSNKPVA-----HVADINSPQGLRWWDSTYANALMA--NGV 117
Qy 177 ALEEKENKILVKETGYFFIYGVLYTDK-----TYAMGHLLIQKKVHVFGDETSVTLFR 231
Db 118 KLE--DNQLVVPADGLYLIYSQVLFPGQCPSTPLFTHTSIRAVS-YQTKVNILSAIK 174
Qy 232 --CIQMPETLP-----NNSCYAGAKLEBEGDELQAIAPRENAQISL-----DGDVTF 278
Db 175 SPCHRETPAEAKPWYPIYGGVQLEKGRDL-----SAEINLPDYLDYAESGVY 227
Qy 279 FGALKL 284
Db 228 FGIAL 233

RESULT 5
I54490
tumor necrosis factor - white-footed mouse
C:Species: Peromyscus leucopus (white-footed mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Feb-1997
C:Accession: I54490
R:Crew, M.D.; Filipowsky, M.E.
Immunogenetics 35, 351-353, 1992
A:Title: Sequence of the tumor necrosis factor/cachectin (TNF) gene from Peromyscus leuc
A:Reference number: I54490; MUID:92218012
A:Accession: I54490
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-235 <RES>
A:Cross-references: GB:M92233; NID:g202506; PID:g202507
C:Genetics:
A:Gene: P1TNF
A:Introns: 62/3; 81/1; 97/1

C:Superfamily: tumor necrosis factor

Query Match 7.4%; Score 107.5; DB 2; Length 235;
Best Local Similarity 22.0%; Pred. No. 0.035;
Matches 54; Conservative 46; Mismatches 84; Indels 61; Gaps 14;
Qy 60 CLTVVSFYQVAALQGLDASLRAELQGHAEKLPAGAGAPKAGLEAPAVTAGLKIFEPP 118
Db 32 CLSLFLLVAGATTFLCLLHFGVIGPQREKFP--NLPPIG-----SMAQTILR----- 82
Qy 119 APGEGNSONSNNKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPLLSFKRGSAL 178
Db 82 -----SSQNSSDK-----PVAHVANHOVDQLEWLSRGANALL-----ANGM 120
Qy 179 EEKENKILVKETGYFFIYGVLYTDK-----TYA-MGHLLQKKVHVFGDELSVTLFRCIQ 234
Db 121 DUKDNLQVLPADGLYLIYSQVLFKGGQSSYVLLTHTVSRFAVS-YEDKNVLLSAIK--S 177
Qy 235 NMPETLPNNS-----CYSAGIAKLEBEGDEL--QLAIPR-----ENAIQLDGDVTF 279
Db 178 PCPKETPGESELKPNWYPIYLGQVQLEKGRDLRAEVNLPKYLDFAESGV-----YF 230
Qy 280 GALKL 284
Db 231 GVIAL 235

RESULT 6
SI2606
tumor necrosis factor alpha precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 08-Sep-1997
C:Accession: SI2606; SI7290; SI8965; I46659
R:Draws, R.T.; Corfee, B.W.; Prestwood, A.K.; McGraw, R.A.
Nucleic Acids Res. 18, 5564, 1990
A:Title: Gene sequence of porcine tumor necrosis factor alpha.
A:Reference number: SI2606; MUID:91016861
A:Accession: SI2606
A:Molecule type: DNA
A:Residues: 1-232 <DRE>
A:Cross-references: EMBL:X54001; NID:g2135; PID:g2136
R:Kuhner, P.; Wuethrich, C.; Peterhans, E.; Pauli, U.
Gene 102, 171-178, 1991
A:Title: The porcine tumor necrosis factor-encoding genes: sequence and comparative a
A:Reference number: SI7289; MUID:91340150
A:Accession: SI7290
A:Molecule type: DNA
A:Residues: 1-232 <KUH>
A:Cross-references: EMBL:X54859; NID:g2132; PID:g2134
A:Note: the authors translated the codon GAG for residue 202 as Gly
R:Choi, C.S.; Mollitor, T.W.; Lin, G.F.; Murtaugh, M.P.
submitted to the EMBL Data Library, January 1991
A:Description: Complete nucleotide sequence of a cDNA encoding porcine tumor necrosis
A:Reference number: SI8965
A:Accession: SI8965
A:Molecule type: mRNA
A:Residues: 1-232 <CHO>
A:Cross-references: EMBL:X57321; NID:g2137; PID:g2138
R:Pauli, U.; Beutler, B.; Peterhans, E.
Gene 81, 185-191, 1989
A:Title: Porcine tumor necrosis factor alpha: Cloning with the polymerase chain react
A:Reference number: I46659; MUID:90034181
A:Accession: I46659
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 44-232 <PAU>
A:Cross-references: GB:M29079; NID:g164694; PID:g164695
C:Genetics:
A:Introns: 62/3; 78/1; 93/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; transmembrane
F.1-77/Domain: signal sequence #status predicted <SIG>

F:78-232/Product: tumor necrosis factor alpha #status predicted <MAT>

```
Query Match      7.4%; Score 107; DB 2; Length 232;
Best Local Similarity 22.0%; Pred. No. 0.038;
Matches 54; Conservative 40; Mismatches 86; Indels 66; Gaps 11;

QY 60 CLTVSVFYQVAALQGLASLRAELQGHAEKLPAGAPKAGLEAPAVTAGLKIFPPA 119
      |||: || || || || || || || || || || || || || || || || || ||
Db 32 CLSFLSFLVAGATLFCLLHFEVIGPQKEFPAGP-----LSI-NPLA 74

QY 120 PEGNSQNSRNRKRAVGPEETVTDCLQLIADSETPTIQKGSYTFVPWLLSFKR---GS 176
      |||: || || || || || || || || || || || || || || || || || ||
Db 75 QGLRSSQTS-----DKPAHVYVANVKAEGQ-----LOWQSGYANALLAN 114

QY 177 ALEPEKENKILVKETGYFFIYQVLYTDK-----TYANGHLIQRKKVHVFGDELSLVTFLR 231
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 115 GVKLKDNQLVVPDGLYLIYSQVLFGRGCGPSTNVLTHTRIAVS-YQTKVNLLSAIK 173

QY 232 --CIQNPETLPNNSCYS----AGIAKLEGEDELQLAIPRENAQISL-----DGDVTF 278
      |||: || || || || || || || || || || || || || || || || || ||
Db 174 SPCQRETPGEAGKWPYEIYLVGVFQLEKDDRL-----SAEINLPDYLDFAESQVY 226

QY 279 FGALKL 284
      |||: || || || || || || || || || || || || || || || || || ||
Db 227 FGIIAL 232

RESULT 7
S06192
tumor necrosis factor alpha precursor - goat (fragment)
N:Alternate names: cachectin; TNF alpha
C:Species: Capra aegagrus hircus (domestic goat)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 08-Sep-1997
C:Accession: S06192; S41867
R:Goldstein, I.M.; Henner, D.; Talhouk, A.
submitted to the EMBL Data Library, March 1989
A:Reference number: S06192
A:Accession: S06192
A:Molecule type: mRNA
A:Residues: 1-193 <GOL>
A:Cross-references: EMBL:X14828; NID:g992; PID:g993
R:Rimstad, E.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41867
A:Accession: S41867
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 36-38,'S','40-78','A','80-88','N','90-114','Q','116-123','D','125-144','G','145-173','L'
A:Cross-references: EMBL:X77317; NID:g452607; PID:g452608
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein

Query Match      7.4%; Score 107; DB 2; Length 193;
Best Local Similarity 23.4%; Pred. No. 0.03;
Matches 50; Conservative 36; Mismatches 92; Indels 36; Gaps 10;

QY 84 QGHAEKLP-----GAGAPKAGLEAPAVTAGLKIFEPAPGEGNSQNSRNRKRAVQGP 140
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 3 RSHHA-LLPALRNRPEE--EQSP---AGPSFNRLVQTLRSSQASNNKPYA----- 52

QY 141 TVTODCLQIADSETPTIQKGSYTFVPWLLSFKRGSALAEKENKILVKETGYFFIYQVQL 200
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 52 -----HVVANISAP-----GQLRWGDSYANALKANGVELKDNQLVVPDGLYLIYSQVL 100

QY 201 Y-----TDKIYAMGHLIQRKKVHVFGDELSLVTFLR--CIQNPETLPN---NSCYAGI 250
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 101 FRHGCGSTPLFTHTRIAVS-YQTKVNLLSAIKSPCHRETPEAEKWPYEIYQGVY 159

QY 251 AKLEGEDELQLAIPRENAQISLSDGDVTFFGALKL 284
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 160 FQLEKGDRLSANEINQPEYLDYAESGQVYFGIIAL 193
```

```
RESULT 8
S13114
tumor necrosis factor alpha - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Sep-1997
C:Accession: S13114
R:Young, A.J.; Hay, J.B.; Chan, J.Y.C.
Nucleic Acids Res. 18, 6723, 1990
A:Title: Primary structure of ovine tumor necrosis factor alpha cDNA.
A:Reference number: S13114; MUID:91067496
A:Accession: S13114
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <YOU>
A:Cross-references: EMBL:X55966; NID:g1403; PID:g1404
C:Superfamily: tumor necrosis factor
C:Keywords: transmembrane protein
```

```
Query Match      7.3%; Score 106; DB 2; Length 233;
Best Local Similarity 21.7%; Pred. No. 0.046;
Matches 54; Conservative 41; Mismatches 88; Indels 66; Gaps 11;
```

```
QY 58 SC-CLTVSVFYQVAALQGLASLRAELQGHAEKLPAGAPKAGLEAPAVTAGLKIFE 116
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 29 SCWCLSLFSFLVAGATLFCLLHFGVIGPQRESPGPS-----FN 70

QY 117 PPAPGEGNSQNSRNRKRAVGPEETVTDCLQLIADSETPTIQKGSYTFVPWLLSFKRGS 176
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 71 RPLVQILRSSQASNNKPV-----AHVVANISAPQLRWGDSYANALMA-----N 115

QY 177 ALEPEKENKILVKETGYFFIYQVLY-----TDKIYAMGHLIQRKKVHVFGDELSLVTFLR 231
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 116 GVELKDNQLVVPDGLYLIYSQVLFGRGCGPSTPLFTHTRIAVS-YQTKVNLLSAIK 174

QY 232 --CIQNPETLPN-----NSCYSAGIAKLEGEDELQLAIPRENAQISL-----DGD 275
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 175 SPCHR---ETLEGAERKWPYEIYQGVFQLEKGDRL-----SAEINLPDYLDYAESG 224

QY 276 VTFPGALKL 284
      |||: || || || || || || || || || || || || || || || || || ||
Db 225 QVYFGIIAL 233
```

```
RESULT 9
JH0529
tumor necrosis factor alpha precursor - sheep
N:Alternate names: cachectin; TNF alpha
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Sep-1997
C:Accession: JH0529; S48118; S20661
R:Green, I.R.; Sargan, D.R.
Gene 109, 203-210, 1991
A:Title: Sequence of the cDNA encoding ovine tumor necrosis factor-alpha: problems wi
A:Reference number: JH0529; MUID:92112044
A:Accession: JH0529
A:Molecule type: mRNA
A:Residues: 1-234 <GRE>
A:Cross-references: EMBL:X55152; NID:g1405; PID:g1406
A:Experimental source: alveolar macrophage
R:Nash, A.D.; Barcham, G.J.; Brandon, M.R.; Andrews, A.E.
Immunol. Cell Biol. 69, 273-283, 1991
A:Title: Molecular cloning, expression and characterization of ovine TNF-alpha.
A:Reference number: S48118
A:Accession: S48118
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-234 <NAS>
A:Cross-references: EMBL:X56756; NID:g297806; PID:g297807
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
```


C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein
F:1-81/Domain: signal sequence #status predicted <SIG>
F:82-234/Product: tumor necrosis factor #status predicted <MAT>

Query Match 6.8%; Score 98; DB 2; Length 234;
Best Local Similarity 21.2%; Pred. No. 0.23;
Matches 49; Conservative 46; Mismatches 72; Indels 64; Gaps 13;
QY 112 LKIFPPAPGEGNSSNSRNKRA-----VGGPEETVQDCL 147
DB 10 VLAEGPLPKAGGPGS--KKCLSLSFLLVAGATILFLLHPRVIGPQEEESPNNL 67
QY 148 QLIAD-SETPTIQKGSYTF-----VPLLKFKGSAL-----EKENK 184
DB 68 HLVNVPVQAVMTLSASRALSDRPLAHVNAVPOVEGQLWL--SQANALLANGMKLTDNQ 125
QY 185 ILVKETGTFYVGVLYTD---KYTA-MGHLIQRKKVHVFGEDELSTVTLFR--CIGNMPE 238
DB 126 LVVPADGLYLYISQVLFSGQGRSYVLLTHTVSREAVS-YPNKNVLLSAIKSPCHRETPE 184
QY 239 TLPNNSCYS---AGTAKLEEGDELQLAIPR-ENAOISLDGDTFFGALKL 284
DB 185 EAPNAPWYEPYILGVGFQLEKGRDRLSTEVNQPEYLDLAEAGQV-YFGIIAL 234

RESULT 13
I38707
Fas ligand - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Aug-1998
C:Accession: I38707; J02340; S57565; I38554
R:Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A:Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A:Reference number: I38707; MUID:95127360
A:Accession: I38707
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-281 <RES>
A:Cross-references: EMBL:U11821; NID:g595430; PID:g595431
R:Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A:Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A:Reference number: J02340
A:Accession: J02340
A:Molecule type: DNA
A:Residues: 1-281 <MIT>
A:Cross-references: GB:D38122; DBJ:D29820; NID:g601892; PID:d1007896; PID:g1369902
R:Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57565
A:Accession: S57565
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-281 <SCH>
A:Cross-references: EMBL:X89102; NID:g887455; PID:g887456
R:Alderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; Go
J. Exp. Med. 181, 71-77, 1995
A:Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A:Reference number: I38554
A:Accession: I38554
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-281 <RES>
A:Cross-references: EMBL:U08137; NID:g624627; PID:g624628
C:Genetics:
A:Gene: FasL
C:Introns: 151/2; 116/3
C:Keywords: glycoprotein; transmembrane protein
F:80-102/Domain: transmembrane #status predicted <TMW>
F:76,184,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7%; Score 97.5; DB 2; Length 281;
Best Local Similarity 20.7%; Pred. No. 0.32;
Matches 56; Conservative 39; Mismatches 107; Indels 69; Gaps 10;
QY 31 LPRKSPSVR---SSKDGKLLAATLLALLSCCLTVVSFYQVAALQGDLSLR-AELOGH 86
DB 63 LPPLPLPLPKRGHSTGLCLVMEFMVLVALVGLGLGMFQLHQLKELAEALRESTSQMH 122
QY 87 HAELKIPAGAGAPKAGLEAPAVTAGLKIEPPAPGEGNSSNSRNKRAVQGPETVTQDC 146
DB 123 TASSLEKQIGHP-----SPP-----PEKELRV 146
QY 147 LOLIADSET---PTIQKGSYTFVPLLSPKRSALKEKENKILVKETGYFFYGVLY-- 202
DB 147 AHLTKSNSRSRPLEWEDYIGV--LL-----SGVYKKGGLVINETGLYFYVSKYFRG 199
QY 202 -----TDKTYAMGHLIQRKKVHVFGEDELSTVTLFRQIONNPETLPNNSCYSAGIAKL 253
DB 200 QSCNNLPLSHKVMRNSKYPQDLVMEGKMSYCTTGQ-----MWARSSYLGAFFNL 251
QY 254 EGGDELQLAIPRENAOISLDGDTFFGALKL 284
DB 252 TSADHLYNVV-SELSLVNFEESQTFEGLYKL 281

RESULT 14
RGBYD2
translation regulator GCD2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G4615; protein YGR083c
C:Species: Saccharomyces cerevisiae
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 12-Dec-1997
C:Accession: S05809; S64378
R:Padon, C.J.; Hannig, E.M.; Hinnebusch, A.G.
Genetics 122, 551-559, 1989
A:Title: Amino acid sequence similarity between GCN3 and GCD2, positive and negative
A:Reference number: S05809; MUID:89339141
A:Accession: S05809
A:Molecule type: DNA
A:Residues: 1-651 <PAD>
A:Cross-references: EMBL:X15658; NID:g3733; PID:g3734
R:Wedler, H.; Scharfe, M.; Wedler, E.; Wambutt, R.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64356
A:Accession: S64378
A:Molecule type: DNA
A:Residues: 1-651 <WED>
A:Cross-references: EMBL:Z72868; NID:g1323118; PID:e243431; MIPS:YGR083
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD;GCD2; GCD12
A:Cross-references: SGD:S0003315; MIPS:YGR083c
A:Map position: 7R
C:Superfamily: translation regulator GCD2
C:Keywords: P-loop; phosphoprotein
F:573-580/Region: nucleotide-binding motif A (P-loop) #status atypical
F:106/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pre
F:121/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status pre

Query Match 6.4%; Score 93.5; DB 1; Length 651;
Best Local Similarity 19.1%; Pred. No. 2;
Matches 67; Conservative 55; Mismatches 140; Indels 89; Gaps 15;
QY 2 DDSTEREQSRLTSLCKREEMKKECVSI--LPRKE-----SPSVRSSRSGKLLAA 50
DB 81 EQQKREQKQKANKKKQNERVKKSTLFGHLETTERRATILATSAVSSPTSITAA 140
QY 51 TLLALLSCCLTVVSFYQVAALQGDLSLR-AELQGHAEKLPAGAGAPKAGLEAPAVTA 110
DB 141 GLMVPVVA-----SALSGSNVLTASSL-----MPVGPNASVTSASAPASIT 182
QY 111 GLKIEPPAPGEGNSSNSRNKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFYFWL- 170

```

Db 183 TTLPASSAALSAGTSSASTNPTAIQ--QETASSNASDVAKTLASISLEAGEFNVIPGIS 240
QY 170 -----LSFKRGSALKEE-KE---NK-----ILVKETGYFFIYG----- 198
Db 241 SVIPTVLEQSFQDNSSLISSSKELLNKLHPSILLTSLAHYKIVGSIPTCIAMLEVF 300
QY 198 QVLYTDKTYAMGHLIQRKVVHVFGEDELSL-----VTLFRCIQNNPETL-----PN 242
Db 301 QIVIKDYQTFKGTTLNLTSLYLSHQIDLLKARPPLSVTMGNAINRLWKQELSLIDPSTPD 360
QY 243 NS-----CYSAG-----IAKLEEGDELOLAIPRENAQISLDGDVTFFGALKLL 285
Db 361 KAAKKDLCEKIGQFAKEKIELAD--QLIIDNASTQIEESTIIVTYGSSKVL 409

RESULT 15
S52715
tumor necrosis factor alpha precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 08-Sep-1997
C:Accession: S52715
R:Meritens, B.; Gaigullis, L.
submitted to the EMBL Data Library, March 1995
A:Description: Cloning and sequence analysis of cDNAs encoding bovine CD40 ligand and bo
A:Reference number: S52715
A:Accession: S52715
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-185 <MER>
A:Cross-references: EMBL:248808; NID:g755701; PID:g755702
C:Superfamily: tumor necrosis factor

Query Match 6.4%; Score 92.5; DB 2; Length 185;
Best Local Similarity 23.0%; Pred. No. 0.52;
Matches 46; Conservative 37; Mismatches 72; Indels 45; Gaps 10;

QY 103 EEAPAVTAGLKIFEPAPGEGNSSONSRRKRAVQGPETVTQDCLQLIADSETPTIQKS 162
Db 13 EQSP---GGPSINSLVQTLRSSQASSNKPVA-----HVVADINSFGQLRW 57
QY 163 YTFVPWLLSFKRGSALEKENKILVKETGYFFIYGQVLYTDK-----TYAMGHLIQRKV 217
Db 58 DSYANALMA--NGVKLE--DNQLVVPADGLYLYSQVLFRCGCGCPSTPLETHTISRIAV 113
QY 218 HVGDELSLVTLFR--CIQNNPETLP-----NNSCYSAGIAKLEEGDELOLAIPRENAQIS 271
Db 114 S-YQTKVNILSAIKSPCHRETPWAAKWPYEPIYQGGVFQLEKGDRL-----SAEIN 165
QY 272 L-----DGDVTFFGALKL 284
Db 166 LPDYLDYAESGGQVIFGIALL 185
```

Search completed: May 13, 1999, 20:59:03
Job time: 2317 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 16:35:35 ; Search time 1016.5 seconds
(without alignments)
1926.852 Million cell updates/sec

Title: US-09-212-270-6
Perfect score: 1284
Sequence: 1 accgctcggaattccggg.....aaaaaaaaaaaaaaaaaaaa 1284

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database : EST.*

- 1: em_est1.*
- 2: em_est2.*
- 3: em_est3.*
- 4: em_est4.*
- 5: em_est5.*
- 6: em_est6.*
- 7: em_est7.*
- 8: em_est8.*
- 9: em_est9.*
- 10: gb_est1.*
- 11: gb_est2.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: gb_est6.*
- 16: gb_est7.*
- 17: gb_est8.*
- 18: gb_est9.*
- 19: gb_est10.*
- 20: gb_est11.*
- 21: gb_est12.*
- 22: gb_est13.*
- 23: gb_est22.*
- 24: gb_est14.*
- 25: gb_est15.*
- 26: gb_est16.*
- 27: gb_est17.*
- 28: gb_est18.*
- 29: gb_est19.*
- 30: gb_est20.*
- 31: gb_est21.*
- 32: em_est10.*
- 33: em_est11.*
- 34: em_est12.*
- 35: em_est13.*
- 36: em_est14.*
- 37: em_est15.*
- 38: em_est16.*
- 39: em_est17.*
- 40: em_est18.*
- 41: em_est19.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	449.2	35.0	456	25	AA682496 zj18e08.s
C 2	403.6	31.4	452	24	AA166695 z08e02.s

C 3	376.4	29.3	378	28	AA906714
C 4	293	22.8	365	10	R16882
C 5	275.2	21.4	377	10	T87299
C 6	257	20.0	272	20	HUM307C08B
7	243.2	18.9	404	30	AI182472
8	210.6	16.4	458	16	AA254047
9	207.6	16.2	383	24	AA422749
10	153.6	12.0	444	30	AI122485
11	130.4	10.2	367	10	R16934
12	52.6	4.1	434	14	T18718
13	50.8	4.0	450	14	T18301
C 14	49	3.8	588	30	AI118282
C 15	49	3.8	515	31	AI195176
C 16	49	3.8	728	31	AI256333
C 17	48	3.7	744	16	AA220873
C 18	48	3.7	370	16	HSU46347
C 19	48	3.7	513	30	AI169669
C 20	47.6	3.7	231	25	AA693444
C 21	47.6	3.7	615	28	AA945620
C 22	47.4	3.7	596	28	AA945321
C 23	47.4	3.7	601	28	AA945586
C 24	47	3.7	528	30	AI169809
C 25	47	3.7	534	30	AI169826
C 26	47	3.7	511	30	AI171464
C 27	46.6	3.6	270	26	AA742997
C 28	46.4	3.6	441	28	AA945161
C 29	46.2	3.6	290	13	N81164
C 30	46	3.6	464	28	AA945600
C 31	46	3.6	425	30	AI169875
C 32	46	3.6	543	30	AI169935
C 33	46	3.6	427	30	AI169981
C 34	45.6	3.6	417	23	AI288116
C 35	45.4	3.5	398	10	T68838
C 36	45.4	3.5	317	29	AA829333
C 37	45.4	3.5	497	30	AI169775
C 38	45.2	3.5	425	31	AI225129
C 39	45	3.5	404	28	AA945000
C 40	45	3.5	544	28	C93698
C 41	45	3.5	569	30	AI170010
C 42	45	3.5	539	30	AI170036
C 43	45	3.5	442	31	AI190896
C 44	44.8	3.5	260	21	AA279795
C 45	44.8	3.5	187	31	AI144047

ALIGNMENTS

RESULT 1
AA682496/C

LOCUS zj18e08.s1

DEFINITION Soares fetal liver spleen INFLS S1 Homo sapiens cDNA

ACCESSION AA682496

NID q2669777

KEYWORDS EST

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 456)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,

Krivan,W., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marr,M.,

Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,

Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.

TITLE WashU-NCI human EST Project

JOURNAL Unpublished (1997)

COMMENT

AA682496 456 bp mRNA EST 19-DEC-1997
zj18e08.s1 Soares fetal liver spleen INFLS S1 Homo sapiens cDNA
clone 450662 3', mRNA sequence.

AA682496
q2669777
EST
human.
Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 456)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krivan,W., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marr,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 456.

FEATURES

source
 1. .456
 Location/Qualifiers
 /organism="Homo sapiens"
 /note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 This is a subtracted version of the original Soares fetal
 liver spleen INFLS library. 1st strand cDNA was primed
 with a Pac I - oligo(dT) primer [5',
 AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="450662"
 /clone_lib="Soares fetal liver spleen INFLS S1"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"

BASE COUNT

137 a 90 c 85 g 144 t

Query Match 35.0%; Score 449.2; DB 25; Length 456;
 Best Local Similarity 99.3%; Pred. No. 3.5e-101;
 Matches 451; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 807 gatettacatttggcttcacgtcttcaaaaggggaagtcgccttagaagaaa 866
 |||||
 Db 456 GATCTTACACATTTGTCATGGCTTCTCAGCTTTAAAGGGGAAGTGCCCTAGAGAAA 397
 |||||
 QY 867 aagagataaatttgtcaagaaactggcttacttttttatataggtcaggttttat 926
 |||||
 Db 396 AAGAGATAAATATTTGGTCAAGAAACTGGTTACTTTTTTATATATGGTCAGGTTTTAT 337
 |||||
 QY 927 atactgaagaccctagccatggacatcaattcagaggaaggaagtcacgtctttg 986
 |||||
 Db 336 ATACTGATAGACCTAGCCATGGGACATCTATTTCAGAGGAGAGAGTCATGTCATTG 277
 |||||
 QY 987 ggaatgaattgctgtgacttggcttcttgcgtgtattcctggaacactac 1046
 |||||
 Db 276 GGGATGAATTGAGTCTGGTGTGCTTGTTCGATGTATTCAAAATATGCCGTGAACACTAC 217
 |||||
 QY 1047 ccaataattcctgctattcagctggcattgcaaaactgggaagagagatgaactccaac 1106
 |||||
 Db 216 CCAATAATTCCTGCTATTCAGCTGGCATTCGAAACATGGAAGAGAGATGAACCTCAAC 157
 |||||
 QY 1107 ttgcaataccaagaagaaatgcacaaatatacactggatggagatgtcacatttttggtg 1166
 |||||
 Db 156 TTGCAATACCAAGAGAAATGCACAAATATCAGTGATGAGATGTCACATTTTGTGCTG 97
 |||||
 QY 1167 cattgaactgctgtgacctacttacaccatgtctgtagttatttctccctttctctg 1226
 |||||
 Db 96 CATTTGAACCTGCTGTGACCTACTTACACCATGCTGTAGTATTTCCTCCCTTCTCTG 37
 |||||
 QY 1227 taccctcaagaagaagaatcctaactgaaataac 1260
 |||||
 Db 36 TACTCGAAGAGAAAGAAATCTAACTGTAATTAC 3

RESULT 2
 AA166695/C
 LOCUS
 DEFINITION z08Se02.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA
 clone 593690 3', mRNA sequence.
 AA166695
 ACCSSION
 NID 91745159

KEYWORDS

human.
 Homo sapiens
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1. (bases 1 to 452)
 Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
 Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
 Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 WashU-NCI human EST Project
 Unpublished (1997)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 533 Std Error: 0.00
 Seq primer: -40m13 fwd. from Amersham.

FEATURES

source
 1. .452
 Location/Qualifiers

/organism="Homo sapiens"
 /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Cloned unidirectionally. Primer: Oligo dT.
 Papillary serous carcinoma, isolated from ascites, 64 year
 old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR
 vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'
 adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "
 /db_xref="taxon:9606"
 /clone="593690"
 /clone_lib="Stratagene ovarian cancer (#937219)"
 /sex="female"
 /dev_stage="adult, 64 years"
 /lab_host="SOLR (kanamycin resistant)"
 /lab_host="89 c 86 g 144 t 2 others

BASE COUNT

131 a 89 c 86 g 144 t

Query Match

Best Local Similarity 97.8%; Pred. No. 6.4e-90;
 Matches 440; Conservative 0; Mismatches 6; Indels 4; Gaps 3;

QY 814 cacattgtccatgctctcagctttaaagggaagtcgccttagaagaaaagaa 873
 |||||
 Db 446 CACATTTGTTCCA-GGCTTCTCAGCTTTAAAGGGGAAGTGCCCTGAAG--AAAAAAGAA 390
 |||||
 QY 874 taaaattattggtcaaaagaactggttacttttttatattatggtcaggttttatatactga 933
 |||||
 Db 389 TAAATATTGGTC-AAGAACTGGTTACTTTTATATATGGTCAGGTTTATATACTGA 331
 |||||
 QY 934 taagacctacgacctggacatctaatcagaggaagaggtccatgtctttgggatga 993
 |||||
 Db 330 TAAGACCTACGCCATGGGACATCAATTCAGAGGNAAGGTCATGCTCTTTGGGATGA 271
 |||||
 QY 994 attgagctggtgacttgttcgatgtattcacaatatccctgaaacacacccaataa 1053
 |||||
 Db 270 ATTGAGTCTGCTGACTTTGTTTCGATGATTCAAATATATGCTTGAACACTACCCCAATAA 211
 |||||
 QY 1054 ttctgtctattcagctggcattgcaaaactgggaagagagatgaactccaacttgcacat 1113
 |||||
 Db 210 TTCCTGCTATTACGTGGCATTCGAAACCTGGAAGAGAGATGAACCTCAACTTGAAT 151
 |||||
 QY 1114 accaagagaaaatgcacaaatatacctggtgagatgtcacatatttttgggtcattgaa 1173
 |||||
 Db 150 ACCAAGAGAAAATGCAAAATATACATGGATGGAGATGTCACATTTTGGTGCATTGAA 91
 |||||
 QY 1174 actgctgtaacctactacacatgctctagctatttctccctcttctcttaacctct 1233
 |||||
 Db 90 ACTGCTGTGACCTACTTACACCATGCTGTAGTATTTTCTCCCTCTCTCTGACTCTCT 31

[illegible]

```

Db 360 CAGGTTTTATANACCTGATAGAGCCCGCCAGCCGACCTTAANTCCAGAGGAGAA 301
QY 973 ggttccatgtcttttgggagtgagtgctggtggaacttggctgttgcgtgtattcacaata 1031
Db 300 GGNCCCATGTCTTTGGGATGAANTGAGTCTGTGTGACTTTGTTTCGATGTATTCAAAATA 241
QY 1032 tgctgaaacactaccacaataatccctgctatttcagctggtgcctggcctgcaaaactggaagaag 1091
Db 240 TGCCCTGAANAACCTACCCCAATTAANTCCCTGCTANTCAGCTGGCATTGCAAAACTGGAAGAAG 181
QY 1092 gagatgaactccaacttgcaatacccaagagagaataatgcacaaatatcactggtgagatg 1151
Db 180 GAGATGAACCTCAACTTGCAATNCCAGAGAAATGACAAATATCAGTGGATGGAGATG 121
QY 1152 tcacatttttggcgactgaaactgctgtgacctaacttaccacatgtctgttagctattt 1211
Db 120 TCACATTTTGGTGCATGTAATGAACTGCTGTGACCTACTTANACCATGTCTGTAGCTATT 61
QY 1212 tctcccttctctgtacctctaagaagaagaatactaaactgaaatacaaaaaaaa 1271
Db 60 TCCCTCCCTTCTGTACTCTAAGAGAAAGAAATCTAACTGAAATACCAAAAAAAA 1

RESULT 5
T87299/c 5
LOCUS HUM307C08B 377 bp mRNA EST 17-MAR-1995
DEFINITION Human aorta cDNA 5'-end GEN-307C08, mRNA sequence.
ACCESSION D79690
NID g1180041
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone_lib:human aorta polyA+.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 377)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisakis,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 145
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1. .377
/organism="Homo sapiens"
/clone="115371"
BASE COUNT 102 a 73 c 72 g 107 t 23 others
ORIGIN

Query Match 21.4%; Score 275.2; DB 10; Length 377;
Best Local Similarity 93.8%; Pred. No. 3.2e-58;

Db 360 CAGGTTTTATANACCTGATAGAGCCCGCCAGCCGACCTTAANTCCAGAGGAGAA 301
QY 919 ggtttatatactgataagacctac-gccatgggacatctaatc-agaggaagaagtc 976
Db 365 GTTTTCATATACTGATAAGACCTACGCCATGGACATCTAATTCNAGAGNAGAAGTC 306
QY 977 catgtcttttgggagatgaattgagt-ctgtggaactttgtttc-gatgtattcacaataatgc 1034
Db 305 CATGTCNTTGGGATGAATTTGAGTCCCTGGTGGTCTTTTCGGATGTATTTCNAAATATGC 246
QY 1035 c-tgaacacactaccacaataatctctgtattcagctggtgcattgcacaaactgggaagaagga 1093
Db 245 CTGGAANAACCTACCCATTAANTCTGCTATTTCAGCTGGCATTGCAAACTGGAAGAAGGA 186
QY 1094 gatgaactccaacttgcaatacccaagagagaataatgcacaaatatcactggtgagatgct 1153
Db 185 GATGAACCTCCNACTTGCATATCCNAGAGCAAAATGCACAAATATCAGCTGGATGGAGATGC 126
QY 1154 acatttttggcgactgaaactgctgtgacctaacttaccacatgtctgttagctattt 1213
Db 125 ACATTTTTCGTGCATTTGAANTGCTGTGACCTACTTANACCATGTCTGTAGCTATTTC 66
QY 1214 ctcccttctctgtacctctaagaagaagaatactaaactgaaatacaaaaaa 1267
Db 65 CTCCCTTCTCTGTACTCTTGAAGANAGAAATCTAACTGNAATAATCANNAAAA 12

RESULT 6
HUM307C08B
LOCUS HUM307C08B 272 bp mRNA EST 14-DEC-1995
DEFINITION Human aorta cDNA 5'-end GEN-307C08, mRNA sequence.
ACCESSION D79690
NID g1180041
KEYWORDS EST; EST(expressed sequence tag); Human aorta.
SOURCE Homo sapiens aorta cDNA to mRNA, clone_lib:human aorta polyA+.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 272)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E., Hirai,Y.,
Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Large-scale sequencing project at Otsuka GEN Research Institute
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 272)
AUTHORS Fujiwara,T.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Tsutomu
Fujiwara, Otsuka GEN Research Institute, Otsuka Pharmaceutical
Co.,Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,
Japan
COMMENT Submitted (7-Nov-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical CO.,Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone:0886-65-2888
Fax :0886-37-1035.
Location/Qualifiers
1. .272
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="human aorta polyA+"
/tissue_type="aorta"
BASE COUNT 83 a 56 c 48 g 79 t 6 others
ORIGIN

```

```

Query Match      20.0%; Score 257; DB 20; Length 272;
Best Local Similarity 97.4%; Pred. No. 9.6e-54;
Matches 265; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Qy 989 gatgaattgagctcgtggaactttgttggatgtattcctgagatgctgctgaaacatacc 1048
|||||
Db 1 GATGAATTGAGCTCGTGGTGCATTTGTTTCGATGTATTCAAAATATGCTTGAACACATACC 60
|||||

Qy 1049 aataattcctcttcagctggtcattgcaaaactggaagagagatgaactcaactt 1108
|||||
Db 61 AATAATTCCTCTATTACCTGCGCATTCGCAAACTGGAAGAGAGATGAATCCCACTT 120
|||||

Qy 1109 gcaataccaagagaaaaatgcacaaatatactggtgagatgtcacattttttgtgtca 1168
|||||
Db 121 GCAATACCAAGAGAAATGCACAAATATACCTGGATGAGATGTCACATTTTGGTGCA 180
|||||

Qy 1169 ttgaaactgctgacactactacacatgctctgttagtattttctccctctctctgta 1228
|||||
Db 181 TTGAAACTGCTGTGACCTACTTACACCATGCTGTGTAGCTATTTCCTCCCTTCTCTGTA 240
|||||

Qy 1229 cctctaagaagaagaatactaaactgaaaaatc 1260
|||||
Db 241 CCYTT-NGRAGRAGAAATCTTAACGTGAAATAC 271
|||||

RESULT 7
A1182472 404 bp mRNA EST 08-OCT-1998
LOCUS uc27f12.1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
DEFINITION 1399247 5', mRNA sequence.
ACCESSION A1182472
NID G3733110
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 404)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:910963
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 375.
Location/Qualifiers
1. 404
/organism="Mus musculus"
/strain="C57BL/6J"
/Note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(GT) primer [5,
TGTTACCAATCTGAAGTGGGAGCGCCGCGCAATGTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
FEATURES
source
/db_xref="taxon:10090"
/clone="1399247"
/clone_lib="Soares mouse mammary gland NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 111 a 101 g 98 t
ORIGIN

Query Match      18.9%; Score 243.2; DB 30; Length 404;
Best Local Similarity 78.4%; Pred. No. 2.5e-50;
Matches 304; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Qy 805 aggatcttacacattgttccattcagcttcagctttaaaggaggagcgcctagaaga 864
|||||
Db 4 AGGAACCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAGAGGAAATGCTTGGAGGA 63
|||||

Qy 865 aaagagaataaaatattggtcaaaactggttacttttttatatatgtgcaggtttt 924
|||||
Db 64 GAAGAGAACAAATAGTGGTGAGGCAAAACAGGCTATTTCATCTACACCGAGTTCT 123
|||||

Qy 925 atatactgataagacacctacgcccattgggacatctaattcagagggaagggtccattgtctt 984
|||||
Db 124 ATACACGGACCCCATCTTTTGTCTATGGGTGATGTCATCCAGAGGAAGAAAGTACACGTCTT 183
|||||

Qy 995 tggggatgaattgagctggtggaactttgttcgatgtattcacaatatgctcgaaacact 1044
|||||
Db 184 TGGGGACGAGCTGAGCCCTGTGACCCCTGTCGATGTATTTCAGATATATGCCAAACACT 243
|||||

Qy 1045 acccaataattcctgctattcagctgcatgcaaaactgcaaaagagagatgaactcca 1104
|||||
Db 244 GCCCAACAATTCCTGCTACTCGGTGGCATCGGAGGTGGGAAGAGAGATGAGATTCA 303
|||||

Qy 1105 acttgcaataccaagagaaaaatgcacaaatatactgagatgagatgtcacattttttgg 1164
|||||
Db 304 GCTTGCAATTCCTCGGGAGATGCACAGATTTTCAGCAACGAGACGACACCTTCTTTGG 363
|||||

Qy 1165 tgcattgaaactgctgtgacactattac 1192
|||||
Db 364 TGC-CTAAAACTGCTGTAACCTCACTTGC 390
|||||

RESULT 8
AA254047 458 bp mRNA EST 14-MAR-1997
LOCUS val10h03.1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
DEFINITION 722549 5', mRNA sequence.
ACCESSION AA254047
NID g1888612
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
AUTHORS Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
1 (bases 1 to 458)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

```


FEATURES
source

Location/Qualifiers
1. .450
/organism="Zea mays"
/strain="W64A2"
/note="Vector: ZipLox; Site.1: SalI; Site.2: NotI; ds-cDNA was prepared from oligo-dr selected mRNA by priming with a NotI oligo- dr oligomer and then adding the second strand to Rnase-nicked DNA:RNA hybrid with DNA PolI. SalI adaptors were added to the ends, the ds-cDNAs were then digested with NotI and size-selected. These were directionally-cloned into the ZipLox phage vector, excised as plasmids, and then individually analyzed."
/db_xref="taxon:4577"
/clone="5c06d10"
/clone_lib="membrane-free polysomes from endosperm"
/lab_host="DH10B"
<1. .>450 119 a 91 c 127 g 110 t 3 others

BASE COUNT
ORIGIN

119 a 91 c 127 g 110 t 3 others

Query Match 4.0%; Score 50.8; DB 14; Length 450;
Best Local Similarity 77.2%; Pred. No. 0.0081;
Matches 61; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 accggtccggaattccgggtgaccacgcggtccgccacgcgtccgagaagacttga 60
|||||
Db 3 ACCGGTCGGGAATTCGGGGTCACCCACCGCTCCGCCACGCGTCGGGGAAACAATC 62
|||||
QY 61 aattctacaaaactgaa 79
|||||
Db 63 TGTGCTGCTCAACGAAA 81
|||||

RESULT 14
AII18282/c

LOCUS AII18282 588 bp mRNA EST 02-SEP-1998
DEFINITION ue39f01.x1 Sugano mouse liver mlia Mus musculus cDNA clone 1482745
3' similar to gb:A06977 SERUM ALBUMIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AII18282
NID 93518606
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 588)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:931101
Seq primer: custom primer used
High quality sequence stop: 393.
Location/Qualifiers
1. .588
/organism="Mus musculus"
/strain="C57BL"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII

FEATURES
source

1. .588
/organism="Mus musculus"
/strain="C57BL"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII

(CACTGTGTG); Site.2: DraIII (CACTATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACTATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGGG and 3' end primer CGACTGCAGTCGAGCACA."
/db_xref="taxon:10090"
/clone="1482745"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
135 a 118 c 134 g 201 t

BASE COUNT
ORIGIN

Query Match 3.8%; Score 49; DB 30; Length 588;
Best Local Similarity 71.9%; Pred. No. 0.023;
Matches 64; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1196 atgtctgtagctattctctccttctctgtaccttaagaagaagaacttaactgaa 1255
|||||
Db 92 ATTCTTTAAACATTGACTTCTGTCTGCTGCTCAATTAATAAATAATGGAAGAAA 33
|||||
QY 1256 atatacaaaaaaaaaaaaaaaaaaaaaa 1284
|||||
Db 32 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4
|||||

RESULT 15
AII95176/c

LOCUS AII95176 515 bp mRNA EST 14-OCT-1998
DEFINITION ui61e11.x1 Sugano mouse liver mlia Mus musculus cDNA clone 1886924
3' similar to gb:A06977 SERUM ALBUMIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AII95176
NID 93747782
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 515)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:971248
Seq primer: custom primer used
High quality sequence stop: 406.
Location/Qualifiers
1. .515
/organism="Mus musculus"
/strain="C57BL"

Search completed: May 13, 1999, 19:44:05
Job time: 11310 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 01:59:41 ; Search time 2254.51 Seconds
(without alignments)
2037.503 Million cell updates/sec

Title: US-09-212-270-6

Perfect score: 1284

Sequence: 1 accggtccggaattcccgagg.....aaaaaaaaaaaaaaaaaaaa 1284

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database :

GenEmbl.*

1: gb_ba1.*

2: gb_ba2.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl1.*

9: gb_pl2.*

10: gb_pri.*

11: gb_pr2.*

12: gb_pr3.*

13: gb_ro.*

14: gb_st.*

15: gb_sy.*

16: gb_un.*

17: gb_vi.*

18: gb_htg.*

19: em_ba.*

20: em_fun.*

21: em_hum1.*

22: em_hum2.*

23: em_in.*

24: em_om.*

25: em_or.*

26: em_ov.*

27: em_pat.*

28: em_ph.*

29: em_pl.*

30: em_ro.*

31: em_sy.*

32: em_un.*

33: em_vi.*

34: em_htg.*

35: em_sts.*

36: gb_ba1.*

37: gb_ba2.*

38: gb_pl1.*

39: gb_pl2.*

40: gb_pri.*

41: gb_pr2.*

42: gb_pr3.*

43: gb_sts.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

C	1	293	22.8	365	43	G30081	G30081 human STS S
C	2	55.6	4.3	7218	6	I66494	I66494 Sequence 14
	3	48.2	3.8	625	3	TSU88240	U88240 Trichinella
	4	44.4	3.5	2078	6	E00096	E00096 DNA coding
	5	44.2	3.4	1720	11	AF030165	AF030165 Homo sapi
	6	44.2	3.4	1720	41	AF030165	AF030165 Homo sapi
	7	44	3.4	2998	4	S80361	S80361 rabC-2 be
	8	43.8	3.4	737	3	TC42A	Y13985 trypanosoma
	9	43.8	3.4	1685	8	OSTA274	X91808 O.sativa mr
	10	43.8	3.4	1885	38	OSTA274	X91808 O.sativa mr
	11	43.6	3.4	599	8	TAEPM1	X68289 T.aestivum
	12	43.6	3.4	599	38	TAEPM1	X68289 T.aestivum
	13	43.4	3.4	1148	8	OSY15009	Y15009 Oryza sativ
	14	43.4	3.4	1148	38	OSY15009	Y15009 Oryza sativ
	15	42.8	3.3	2923	8	ATHSP91	Z70314 A.thaliana
	16	42.8	3.3	1577	8	OSCHSLKGN	X91811 O.sativa mr
	17	42.8	3.3	2923	38	ATHSP91	Z70314 A.thaliana
	18	42.8	3.3	1577	38	OSCHSLKGN	X91811 O.sativa mr
C	19	42.6	3.3	116884	18	PFMAL1P1	AL031744 Plasmodiu
C	20	42.4	3.3	202	4	DOGSND17B	M73046 Dog inserte
C	21	42.4	3.3	7192	13	MMAJ2970	AJ222970 Mus muscu
	22	42.2	3.3	1953	4	OAF6DGH	X60195 Ovine 6-pgd
	23	42	3.3	2055	8	TAU48227	U48227 Triticum ae
	24	42	3.3	1056	13	MUSUNKNE	L04849 Mouse (clon
	25	42	3.3	1262	13	MUSUNKNE	L04852 Mouse (clon
	26	42	3.3	2055	38	TAU48227	U48227 Triticum ae
	27	41.8	3.3	1566	9	CRO7333	AJ007333 Catharant
	28	41.8	3.3	45892	11	HS104M094	Z98948 Human DNA s
	29	41.8	3.3	1566	39	CRO7333	AJ007333 Catharant
	30	41.8	3.3	45692	41	HS104M094	Z98948 Human DNA s
	31	41.6	3.2	12029	3	AE001389	AE001389 Plasmodiu
	32	41.6	3.2	1725	8	AF007785	AF007785 Zea mays
	33	41.6	3.2	1725	38	AF007785	AF007785 Zea mays
	34	41.4	3.2	1014	6	A30330	A30330 Artificial
C	35	41.4	3.2	1016	6	A30331	A30331 Artificial
	36	41.4	3.2	1013	6	A32826	A32826 Synthetic c
C	37	41.4	3.2	1015	6	A32827	A32827 Synthetic c
	38	41.4	3.2	688	6	I03321	I03321 Sequence 12
	39	41.4	3.2	1013	6	I11571	I11571 Sequence 23
	40	41.4	3.2	1013	6	I24003	I24003 Sequence 30
	41	41.4	3.2	1319	9	AF031609	AF031609 Oryza sat
	42	41.4	3.2	1414	9	ATH7587	AJ007587 Arabidops
	43	41.4	3.2	3832	11	HSU75370	U75370 Human mitoc
	44	41.4	3.2	1319	39	AF031609	AF031609 Oryza sat
	45	41.4	3.2	1414	39	ATH7587	AJ007587 Arabidops

ALIGNMENTS

RESULT	1	G30081/c	human STS SHGC-36171, sequence tagged site.	DNA	STS	04-OCT-1996
LOCUS		G30081	human STS SHGC-36171, sequence tagged site.			
DEFINITION		G30081	human STS SHGC-36171, sequence tagged site.			
ACCESSION		G30081	human STS SHGC-36171, sequence tagged site.			
NID		g1593632	STS; STS sequence; primer; sequence tagged site.			
KEYWORDS		human.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		1 (bases 1 to 365)				
AUTHORS		Myers,R.M.				
JOURNAL		Unpublished (1996)				
COMMENT						

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: GTACAGAGAAAGGAGGAAAAATAGC
Primer B: ATGCACAAATATCACTGGATGG
STS size: 105
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
each 1 uM
Primer:
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from R16802
-- Washington University/Merck EST sequence.

FEATURES

source
1. .365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="13"
43. .147

primer_bind 43. .67

primer_bind 94 a 70 c 76 g 116 t 9 others
BASE COUNT

ORIGIN

Query Match 22.8%; Score 293; DB 43; Length 365;
Best Local Similarity 94.7%; Pred. No. 1.le-62;
Matches 341; Conservative 0; Mismatches 14; Indels 5; Gaps 4;
Qy 917 caggtttatata-ctgataagacctagc-cctggagacatc--taattcagagagaag 972
Db 360 CAGGTTTATANACTGATAGCCGCCACCCGCGCATCTTAANTCCAGAGAGAA 301
Qy 973 ggt-ccatgtcttggggatgaattgagctgtggtgactttgttcgatgtattcaaaata 1031
Db 300 GGNCCCATGCTTTGGGATGAANTGAGTCTGTGACCTTTGTTTCGATGTATCAAAATA 241
Qy 1032 tgcctgaacactaccacaataattcctgctattcagctggcattgcaaaactggaagaag 1091
Db 240 TGCCCTGAANACTACCCAATAANTCCTGCTANTCAGCTGGCATTGCAAAACTGGAAGAAG 181
Qy 1092 gagatgaactccaacttgcaatacaagaagaaatgcacaaatcatcactgagatgagatg 1151
Db 180 GAGATGAATCACTTGCATNCCAGAGAGAAATGACAAATATCACTGGATGGAGATG 121
Qy 1152 tcacatttttgggtgactgaaactgctgacacctacttacacatgctgtagctattt 1211
Db 120 TCACATTTTGGTGATGAAACTGTGTGACCTACTTANACCATGCTCTAGCTATT 61
Qy 1212 tcctcccttctctgaccttaagaagaagaatctaaactgaataaatacaaaaaa 1271
Db 60 TCCTCCCTTCTCTGTACCTTAAGAAGAAAGATCTAACTGAAATACCAAAAAA 1

RESULT 2
I66494/c
LOCUS
DEFINITION
Accession
NID

I66494 7218 bp DNA PAT 23-DEC-1997
Sequence 14 from patent US 5670367.
I66494
g2724471

KEYWORDS
SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 7218)

Unclassified.
Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.

AUTHORS
TITLE
Recombinant fowipox virus

JOURNAL
Patent: US 5670367-A 14 23-SEP-1997;

FEATURES
Location/Qualifiers

source
1. .7218

BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others

ORIGIN

Query Match 4.3%; Score 55.6; DB 6; Length 7218;

Best Local Similarity 2.7%; Pred. No. 0.0011;

Matches 10; Conservative 221; Mismatches 145; Indels 0; Gaps 0;

Qy 48 gagaagactttgaaattcttcaaaaactgaaatgaaatgagaaagacattgagcaa 107
Db 1426 RRR 1367

Qy 108 tccaatcgagggttaaatgccagcaaacctactgtacagttaggggttagagatgcagagaag 167
Db 1366 RRR 1307

Qy 168 gcagaagagagaaattcaggataactctctgaggggtgagcccaagcctgccatgag 227
Db 1306 RRR 1247

Qy 228 tgcacgcagacatcaacaacagacagataacaggaatgatccattccctgtgtcactt 287
Db 1246 RRR 1187

Qy 288 attctaaagcccaactcaagttcaagtagtgatggatgactcacagagaagggg 347
Db 1186 RRR 1127

Qy 348 agcagtcacgccttactcttgccttaagaagaagagaaatgaaactgaagagtggtg 407
Db 1126 RRR 1067

Qy 408 ttctcatctctccacac 423

Db 1066 ATCGCAAGCTCCCTCG 1051

RESULT 3

TSU88240

LOCUS

DEFINITION

Trichinella spiralis

hypothetical ORF 2.20 mRNA, partial cds.

Accession

U88240

NID

g1857757

KEYWORDS

SOURCE

ORGANISM

Trichinella spiralis.

Trichinella spiralis

Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;

Adenophorea; Enoplia; Trichocephalida; Trichocephalatinia;

Trichinelloides; Trichinellidae; Trichinella.

REFERENCE

1 (bases 1 to 625)

Polvere, R.I. and Despommer, D.D.

AUTHORS

Direct Submission

TITLE

Submitted (04-FEB-1997) Environmental Science, Columbia University,

JOURNAL

630 W. 168th Street, New York, NY 10032, USA

FEATURES

source

1. .625

/organism="Trichinella spiralis"

/db_xref="taxon:6334"

/clone="2.20"

<1. .507

/note="hypothetical ORF 2.20"

/codon_start=1

CDS

JOURNAL MEDLINE	Mol. Biochem. Parasitol. 92 (2), 313-324 (1998)
REFERENCE	98319417
AUTHORS	2 (bases 1 to 737)
TITLE	Lopez, M.C. Direct Submission
JOURNAL	Submitted (23-JUN-1997) Lopez M.C., Molecular Biology, Instituto de Parasitologia y Biomedicina 'Lopez Neyra', C.S.I.C., Ventanilla, 11 18001 Granada SPAIN
FEATURES	Location/Qualifiers

```

FEATURES             Location/Qualifiers
     1..737
         /organism="Trypanosoma cruzi"
         /strain="y"
         /db_xref="taxon:5693"
         /dev_stage="Amastigote"
     98..643
         /gene="h2a"
     98..505
         /gene="h2a"
         /codon_start=1
         /evidence=experimental
         /product="Histone H2A"
         /db_xref="PID:e324188"
         /db_xref="PID:g2222802"
         /db_xref="SP:TRMBL:O02617"
         /translation="MATPQAAKAAKKHGGRSAKAGLIFPVGRVGSLLRRQGYARR
IGASGVYAAVALELYTAEELLESVKAASQAQAKPKRLTPRTVTTLAVRHDDDLGMLLK
NTVLSRGVYIPSLNKAVAKKHKSRAATPSA"

```

polyA_site
643
/gene="h2a"
/avidance-overl-mortal

BASE COUNT	219 a	179 c	208 q	131 t
BASE COUNT	219 a	179 c	208 q	131 t

Query Match	3.4%;	Score 43.8;	DB 3;	Length 737;
Best Local Similarity	69.0%;	Pred. No. 0.76;		

QY	1198	gtctgtagctatttttctctccctttctctgtaccttaagaagaagaatactcaactgaaaa	1257
Db	616	gtttgtgtgtgtttttcttttttcgtttttctaaaaaataaaaaaataaaaaaataaaaaa	675

Qy	1258	tacaaaaaaaaaaaaaaaaaaaaaa	1284
Db	676	AAAAAAAAAAAAAAAAAAAAAAAAA	702

RESULT	9
OSTA274	
LOCUS	OSTA274
DEFINITION	O. sativa mRNA for alpha-tubulin
ACCESSION	X91808
NID	g1136123
KEYWORDS	alpha-tubulin; tuba gene.
SOURCE	rice.
	PLN
	22-DEC-1995

Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE

1 (bases 1 to 1685)
Maestroni, A., Giani, S. and Breviario, D.
Rice alpha-tubulin cDNAs
Unpublished
2 (bases 1 to 1685)

Journal: *Journal of Molecular Evolution*
 Submitted (25-SEP-1995) D. Breviario, Istituto Biosintesi Vegetali
 CNR, Via Bassini n 15, I-20133 Milano, ITALY
 COMMENT: Overlaps with Z11931.
 FEATURES: Location/Qualifiers
 source 1..1685
 /organism="Oryza sativa"
 /cultivar="Arborio"

```

/db_xref="taxon:4530"
/dev_stage="3 days old seedling"
/tissue_type="coleoptile"
/clone_lib="lambda zapII"
/clone="OSTA-274"
110..1462
/gene="tuba"
110..1462
/gene="tuba"
/codon_start=1
/product="alfa-tubulin"
/db_xref="PID:e020586"
/db_xref="PID:g1136124"
/translation="MREIISIHIGQAGIQGNACWELYCLEHGIEPTDGTMPSTSYGV
AHDANFTFSETGAGKHVPRAIFVDLEPTVIDEVRTGSYQLFHPPEOLISKGDAAANN
FARGHTVGEIYDCLDRVRKLDNCTGLQGLFVENVAGGCTGSGLGSLLLRLSLVD
YGRKSLGFTIYPSQVSTAVPEPYNYSVTHSLLEHTDVAVLDDNEAIIYDICRRSLD
IERPTNLNRLISQIISSTLSRFDGAINVDVTEFTNLVPYPIRHEFLMSLYAPVI
SABKAYHQSLVPEITNAVPEPSMMAKCDPRHGKYMACCLMTRGDVVPYKDVNAAVAT
IKTKRIVQWDMCPTGFKCGINTQPPSVVPGGLAKVQRAVCMISNNATAVEVFSRID
HKEDLMYAKRAFVHWIYVGEEMSEGSFEAREDLAALEKDYEEVGEAGADDNDDGEDY
1

```

BASE COUNT	426 a	402 c	445 g	412 t
ORIGIN				

Query Match	3.4%;	Score 43.8;	DB 8;	Length 1685;
Best Local Similarity	64.1%;	Pred. No. 0.82;		

Qy	1182	gacctaacttaaccatgtctgtagctatttttctcccttttctctgtacctctaaagaagaa	1241
Db	1582	GCCCCCTGTTGGCCATGATTGTTCATATCTTCCCATTTTGGTGCAAAAAAATAAAAAA	1641

Qy	1242	agaaatctcaactgaaaaatac	aaaaaaaaaaaaaaaaaaaaaaa	1284
Dd	1642	AAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAA	1684

RESULT	10	
OSTA274		
LOCUS	OSTA274	1685 bp RNA PLN
DEFINITION	O.sativa mRNA for alpha-tubulin (clone OSTA-274).	22-DEC-1995
ACCESSION	X91808	
NID	g1136123	
KEYWORDS	alpha-tubulin; tuba gene.	
SOURCE	rice.	

ORGANISM	<i>Oryza sativa</i>
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; <i>Oryza</i> .
AUTHORS	1 (bases 1 to 1685)
TITLE	Maestroni, A., Giani, S. and Breviario, D.
JOURNAL	Rice alpha-tubulin cDNAs
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 1685)
TITLE	Breviario, D.
JOURNAL	Direct Submission
COMMENT	Submitted (25-SEP-1995) D. Breviario, Istituto Biosintesi Vegetali CNR, Via Bassini n 15, I-20133 Milano, ITALY
FEATURES	Overlaps with Z11931. Location/Qualifiers 1..1685

```

/organism="Oryza sativa"
/cultivar="Arborio"
/db_xref="taxon:4530"
/dev_stage="3 days old seedling"
/tissue_type="coleoptile"
/clone_lib="lambda zapII"
/clone="OSTA-274"
110..1462
gene
110..1462
/gene="tubA"
110..1462
CDS

```


[illegible]

[illegible]

```

Db 114 RGETAVTLAEKLAGLRAGGYAKPIVSVEVSVNYITVLQVITAGLQPVDRGVHS 173
QY 90 KLPAGACAPKAGLEEAPAVTAGLKIFEPAPGEGNSQNSRNRKRAVOGPEETVTQCLQL 149
Db 174 EIIARAGGLURADAFVLTFR-----ADGTSKLNKQLAQGGPE----- 214
QY 150 IADSETPTIQKSYTFVPWLLSFKRSALKEENKILVKETGYFFIYGQV-----LVT 202
Db 214 ----QPVVTPGDKLFP-----EVEFYIYGQVNPAGVYAIRT 248
QY 203 DKT-----YAMGHLI-----QRKKVHVFQDEL 226
Db 249 DMTLRRALAQGGGLTPAGSSKRVKVSQDQGEIKL 282

RESULT 9
O77510
ID O77510 PRELIMINARY; PRT; 233 AA.
AC O77510;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA PRECURSOR.
OS PAPIO HAMADRYAS URSINUS (CHACMA BABOON).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; PAPIO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98147379.
RA HAUDEK S.B., REDL H., SCHLAG G., GIROIR B.P.;
RT "Complementary DNA (cDNA) sequence of baboon tumor necrosis factor
alpha."
RL MOL. IMMUNOL. 34:1041-1042(1997).
DR EMBL; AF019963; G3417555; -.
DR PROSITE; PS00251; TNF_1; 1.
KW SIGNAL.
FT CHAIN 1 76 POTENTIAL.
FT CHAIN 77 233 TUMOR NECROSIS FACTOR ALPHA.
SQ SEQUENCE 233 AA; 25658 MW; E38CF6E8 CRC32;

```

```

Query Match 6.0%; Score 87.5; DB 4; Length 233;
Best Local Similarity 19.9%; Pred. No. 5.5;
Matches 49; Conservative 38; Mismatches 92; Indels 67; Gaps 9;

QY 61 LTVSVFYQVAALQGDLSLRAELQGHAEKLPAGACAPKAGLEEAPAVTAGLKIFEPAP 120
Db 33 LSLFSLVAGATTLFCLLHFGVIGPQREFPKDP-----SLISPLAQ 75
QY 121 GEGNSSQNSRNRKRAVOGPEETVTQCLQLIADSETPTIQKSYTFVPWLLSFKRSAL-- 179
Db 76 AVRSSTSDPKPVHVANPQAEGLQ-----WL--NRRANALLA 114
QY 179 ---EERENKILVKETGYFFIYGQVLTDTKYAMGHLIQRKKVH---VFGDELSLVTFLR 231
Db 115 NGVELTDNLVLPSEGLIYQVLFKGGCCPSNVLHLLTHTSIRIAVSYQTKVNLUSAIK 174
QY 232 --CIGNMPETLPNNSCYS-----AGIAKLEEGDELQAIAPRENAQISL-----DGDVTF 278
Db 175 SPQRETPGEAEAKPWYETIYLGGVFQLEKGRDL-----SAEINLPDYLDFAESGQVY 227
QY 279 FGALKL 284
Db 228 FGIALL 233

RESULT 10
O87015
ID O87015 PRELIMINARY; PRT; 927 AA.
AC O87015;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)

```

```

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FIMV.
GN FIMV.
OS PSEUDOMONAS AERUGINOSA.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC PSEUDOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PAOLI;
RA SEMMLER A., WHITCHURCH C.B., MATTICK J.S.;
RT "Pseudomonas aeruginosa twitching motility gene fimv.";
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U93274; G3237312; -.
SQ SEQUENCE 927 AA; 98172 MW; 2642488E CRC32;

Query Match 6.0%; Score 87.5; DB 9; Length 927;
Best Local Similarity 24.1%; Pred. No. 32;
Matches 51; Conservative 24; Mismatches 80; Indels 57; Gaps 8;

QY 3 DSTEREQSLTCLKKREEM--KLKECVSILPRKESPSVRSKDGKLAATLLALLSCC 60
Db 327 DSTRRENEELQSRMODLSQDLKQKLIQL-----KDA----- 360
QY 61 LTVSVFYQVAALQGDLSLRAELQGHAEKLPAGACAPKAGLEEAPAVTAGLKIFEPAP 104
Db 360 -----QLAKLOGQLGAEQGAAPNAAALPDASOPNAAQAAPQPGTAAAAPTAPAG 412
QY 104 EAPAVTAGLKIFEPAP--CEGNSQNSRNRKRAVOGPEETVTQCLQLIADSETPTIQKS 162
Db 413 EAPAPAPQPPVAPPAPAEKPPAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPV 472
QY 163 ---YTFVWLLSFKRSALKEEN--KILVKETG 191
Db 473 ALLALIVLLMILSRNAQKEEAQAFADTG 504

RESULT 11
O54693
ID O54693 PRELIMINARY; PRT; 377 AA.
AC O54693;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE EDA PROTEIN HOMOLOG.
GN TABBY.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97449184.
RA FERGUSON B.M., BROCKDORFF N., FORMSTONE E., NGUYEN T., KRONMILLER J.E.,
RA ZONANA J.;
RT "Cloning of tabby, the murine homolog of the human EDA gene: evidence
for a membrane-associated protein with a short collagenous domain.";
RL HUM. MOL. GENET. 6:1589-1594(1997).
RN [2]
RP SEQUENCE OF 1-132 FROM N.A.
RC STRAIN-129;
RA BROCKDORFF N., ZONANA J., FORMSTONE E., FERGUSON B.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF004435; G2660700; -.
DR EMBL; AF004434; G2660698; -.
SQ SEQUENCE 377 AA; 40002 MW; AEA7ED8 CRC32;

```

```

Query Match 6.0%; Score 87.5; DB 10; Length 377;
Best Local Similarity 18.0%; Pred. No. 10;
Matches 39; Conservative 40; Mismatches 79; Indels 59; Gaps 9;

QY 92 PAGACAPKA--GLEEAPAVTAGLKIFEPAPGEGNSQNSRNRKRAVOGPE-----EETVTQD 145
Db 92 PAGACAPKA--GLEEAPAVTAGLKIFEPAPGEGNSQNSRNRKRAVOGPE-----EETVTQD 145

```

```
Db 190 PGPQPPGPPGIPGIPGI-PGTTVMGPCCP-----PGPPGPPGPPGLOGPSGAADKTGTRE 245
QY 146 CLOLIADSETPTIQKSYTFVPWLLSKFKRGSALEREKENILVKETGYFFIYGQVLYTKT 205
Db 246 NQPAVVHLQ-----GQSAIQVND-----LSGGVLDWSRITNPKV 283
QY 206 YAMGHLIQKRVHVFGEISLVTL-----FCIQNMPTPLN-NSCYS 247
Db 284 F-----KLHPRSGELEVYINFTFASYEVVVYDEKPFQLQCTRSIETGKNTNYCT 334
QY 248 AGIAKEEGDELOLAIPRENAQISLDGDTFFGALKL 284
Db 335 AGVCLKARQKIAVMKHADISINMSKHTFFGAIRL 371

RESULT 12
O70332
ID O70332 PRELIMINARY; PRT; 216 AA.
AC O70332;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR-ALPHA (FRAGMENT).
OS MESOCRICETUS AURATUS (GOLDEN HAMSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; CRICETINAE; MESOCRICETUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN.
RX MEDLINE: 98234044.
RA MELBY P.C., TRYON V.V., CHANDRASEKAR B., FREEMAN G.L.;
RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and
RT analysis of cytokine mRNA expression in experimental visceral
RT leishmaniasis."
RL INFECT. IMMUN. 66:2135-2142(1998).
DR EMBL: AF046215; G3005109; -.
DR PROSITE: PS00251; TNF_1; 1.
FT NON_TER 1
FT NON_TER 216
SQ SEQUENCE 216 AA; 23793 MW; 79A08367 CRC32;

Query Match 6.0%; Score 87; DB 10; Length 216;
Best Local Similarity 21.4%; Pred. No. 5.5;
Matches 34; Conservative 35; Mismatches 56; Indels 34; Gaps 8;

QY 124 NSSQNRKRAVQGPETVTQDCLQIADSETPTIQKSYTFVPWLLSKFKRGSALEREKEN 183
Db 72 SSSQNSDKPVGHVAVHQQVEQLEWLSHRANALLANG-----MSLKDN 115
QY 184 KILVKETGYFFIYGQVLYTKD---TYA-MGHLIQKRVHVFGEISLVTLFR--CIQNP 237
Db 116 QLVIPADGLYLVISQVLFNGQCPVSLVLTHTVSRIASV-YEDVNLLSAIKSPCKETP 174
QY 238 ---ETLP-NNSCYSAGIAKEEGDELOLAIPRENAQISL 272
Db 175 EGEELKPWEPIYLGVFQLEKGDRL-----SAEVNL 206

RESULT 13
O49749
ID O49749 PRELIMINARY; PRT; 1229 AA.
AC O49749;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE P-GLYCOPROTEIN-LIKE PROTEIN.
GN PGP3.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; CHAROPHYTA/EMBRYOPHYTA GROUP; EMBRYOPHYTA;
OC TRACHEOPHYTA; EUPHYLOPHYTES; SPEERMATOPHYTA; MAGNOLIOPHYTA;
OC EUDICOTYLEDONS; ROSIDAE; CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA SIDLER M., DUDLER R.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Y15990; E1217969; -.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-BINDING; TRANSPORT.
SQ SEQUENCE 1229 AA; 133234 MW; 73BDF595 CRC32;

Query Match 6.0%; Score 86.5; DB 8; Length 1229;
Best Local Similarity 22.3%; Pred. No. 56;
Matches 73; Conservative 51; Mismatches 115; Indels 89; Gaps 15;

QY 16 LAKREEMKLECVSILPRKESPVSSKDKLLAALLLALLSCCLTVVSYFYQVAALQ-G 74
Db 502 LKDPRILLDDEATSALDAESERVQALDRVMSRT-----TVIVAHRLSTVRNA 551
QY 75 DIASL--RAEL--QGHAEKLP--AGAGAPKAGLEAPAVTAGLKTFEPAPGEGNSS-- 127
Db 552 DMIAVTHRGKIVVEGSHSELLKDHGAYALRLQ-----KIKKPKRLESSNELR 602
QY 127 -----QNSNRKRAVQGPETVTQDCLQIADSETPTIQKSYTFVPWLLSKFKRGSAL 180
Db 603 DRSINRGSSRNIRTRYHDDDSVS--VLGLGRQENTEISREQSRNV-----SITRIAALNK 656
QY 181 KENKILVKET-----GYEF-IYG-----Q 198
Db 657 PETTILILGLGAVNGTTFPIFGILFAKVIKFAFPKPHDMKRDSEFWMFVLLGVASL 716
QY 199 VLYTKTY----AMGHLIQKRVHVFGEISL-VILFRCIQNMPTPLNNSCYSAGIAKL 253
Db 717 IYVPMHTYLFVAGGRLQIRVMCFKVVHMEVGVDFDPENSSGTIGSRLSADAALIKT 776
QY 254 EGGDELOLAIPRENAQISLDGDTFFGA 281
Db 777 LVGDSLSLSV--KNAAAVSGLIIFTA 802

RESULT 14
O35734
ID O35734 PRELIMINARY; PRT; 233 AA.
AC O35734;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA PRECURSOR.
GN TNF-ALPHA OR TNF ALPHA.
OS MARMOTA MONAX (WOODCHUCK).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; SCIURIDAE; SCIURINAE; MARMOTA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PBMC;
RX MEDLINE: 98139533.
RA LOHRENGEL B., LU M., ROGGENDORF M.;
RT "Molecular cloning of the woodchuck cytokines: TNF-alpha, IFN-gamma,
RT and IL-6."
RL IMMUNOGENETICS 47:332-335(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PERIPHERAL BLOOD;
RA ZHOU H., HU J., SEEGER C.;
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Y14137; E348344; -.
DR EMBL: AF082491; G3450964; -.
DR PROSITE: PS00251; TNF_1; 1.
DR PFAM: PF00229; TNF; 1.
KW SIGNAL.
FT SIGNAL 1 77
FT CHAIN 78 233
FT CHAIN 78 233
SQ SEQUENCE 233 AA; 25764 MW; 6FC0F34A CRC32;
```

Query Match 6.0%; Score 86.5; DB 10; Length 233;
Best Local Similarity 21.2%; Pred. No. 6.6;
Matches 55; Conservative 39; Mismatches 73; Indels 93; Gaps 14;

```
QY 60 CLTVSYQVAALQGDILASRAELQGHAEK----LPAGAGAPKAGLEAPAVTAGLKIF 115
   |||: || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 32 CLSLFSLVAGATFLCFLHFGVIGPQREFLNNFLPLSPQAQMLTLR----- 80
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 116 EPPAPGEGNSSNRKRAVGPPEETV---TODCLOLIADSETPTIOKGSYTFVPWLLSF 172
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 -----SSQNNDK-----PVAHVAKNDEQLV-----WL--S 107
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 173 KRGSAL-----EKENKILVKETGYFFIYQVLYTDK---TYA-MGHLIQKKVHVFGDE 223
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 108 RRANALLANGMELIDNQLVVPANGLYLVISQVLFKGGCPSYVLLTHTVSRFAVS-YQDK 166
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 224 LSLVTLFRCTONMPETLPNNS-----CYSAGIAKLEGGDELQLAIPRENAQIS 271
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 167 VNLSAIAK-----SPCPKESLEGAEPKWPVEPIYLGGVFELQKGDRL-----SAENV 213
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 272 L-----DGDVTFFGALKL 284
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 214 LPSYLDFAESGQVIFGVIAL 233
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 15

```
P90611
ID P90611 PRELIMINARY; PRT; 231 AA.
AC P90611;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE H4 GENE.
GN H4.
OS TOXOPLASMA GONDII.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; COCCIDIA; EIMERIIDA; SARCOCYSTITDAE;
OC TOXOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BK;
RA NCKEMANN S.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y09782; E284082; -.
SQ SEQUENCE 231 AA; 25984 MW; CC32AB49 CRC32;
```

Query Match 5.9%; Score 85.5; DB 3; Length 231;
Best Local Similarity 26.7%; Pred. No. 7.9;
Matches 50; Conservative 21; Mismatches 79; Indels 37; Gaps 8;

```
QY 18 KREEMKLEKCVSILPRKESPSVRSKDGKLLAATLLALLSCCLTVYSFYQVAALQGDLA 77
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43 KREPFLYSRVTML-----RPTVRS-----LJSLGLTVILY---LALTGSAD 82
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 78 SLRAELQGHAE-----KLPAAGAGAPKAGLEAPAVTAGLKIFPPAPGEGNSSNRK 133
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 ALASHLSRHRMEAGRRYMDTNDVESAGRQSEPMADRAQAEHFCAPTQ--SEMKEFOEE 140
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 134 AVQGPETVTODCLOLIADSTP-----TIQKGSYTFVPWLLSFKRGSALAEKENKI-LVK 188
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 IKGVEETKHED-----DPEMTRLMVTEKQESKNFKNMAKSQSFSRIEELGGSISFLT 194
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 189 ETGYFFI 195
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 ETGVMTI 201
   |||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 14, 1999, 21:39:56
Job time: 7990 sec

44	39.4	3.1	2023	1	V59003
45	39.4	3.1	2346	1	V59693

ALIGNMENTS

RESULT 1
V30934
ID V30934 standard; DNA: 1100 BP:

AC V30358, 11-SEP-1998 (first entry)
 DE Homo sapiens neutrokin alpha protein gene.
 DE neutrokin alpha; cell proliferation; differentiation; migration;
 KW cytotoxicity; cell death; treatment; tumour; infection; inflammation;
 KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
 KW fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
 KW cachexia; detection; diagnosis; drug screening; ss.

def; haematopoiesis; sepsis; deficiency syndrome; rheum

	Location/Qualifiers
OS	Homo sapiens.
FH	Key
FT	147..1004
CDS	
KW	cachexia; detection; diagnosis; drug
OS	Homo sapiens.

FR	key	location/qualifiers
FT	CDS	147..1004
FT		/*tag= a
FT		/product= neurokinine alpha
FT		
PN		W09818921-A1.
PD		07-MAY-1998.
PF		25-OCT-1996; U17957.
PR		23-OCT-1996; WO-U17957.
PA		(HUMA-) HUMAN GENOME SCI INC.
PI		Ebner R, N1 J, Yu G;
PI		WPI: 98-272216/24.
DR		P-PSDB; W58391.
DR		
PT		New isolated human Neurokinine alpha - used to develop products for
PT		diagnosis and treatment of e.g. tumours, infections,
PT		immunodeficiencies or autoimmune diseases
PS		Claim 2: Fig 1: 104pp; English.

25-OCT-1996: 01/1957.
 25-OCT-1996: WO-U17957.
 (HUMA-) HUMAN GENOME SCI INC.
 Ebner R, Ni J, Yu G;
 WPI: 98-272216/24.
 P-PSDB: W58391.
 New isolated human Neurotrophin alpha - used to develop products for
 diagnosis and treatment of e.g. tumours, infections,
 immunodeficiencies or autoimmune diseases
 Claim 2; Figure 1: 104pp; English.
 The sequence is that encoding neurotrophin alpha.
 Neurotrophin alpha (NA) polypeptides modulate cell proliferation,
 differentiation, migration cytotoxicity and cell death.
 They can be used to treat e.g. tumour and tumour metastasis, infections
 by bacteria, viruses and other parasites, immunodeficiencies, graft
 inflammatory diseases, lymphadenopathy, autoimmune diseases, graft
 versus host disease and to stimulate peripheral tolerance, destroy some
 transformed cell lines, mediate cell activation and proliferation, and
 are functionally linked as primary mediators of immune regulation and
 inflammatory responses. Such activity is useful for immune enhancement
 or suppression, myeloprotection, stem cell mobilisation, acute and
 chronic inflammatory control and treatment of leukaemia. They can also
 be used to stimulate wound healing and to treat fibrotic disorders
 including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They
 can also be used to regulate haematopoiesis, by regulating the activation
 and differentiation of various haematopoietic progenitor cells, e.g. to
 release mature leukocytes from the bone marrow following chemotherapy.
 NA and in stem cell mobilisation. NA may also be used to treat sepsis.
 NA antagonists can be used to prevent septic shock, inflammation, cerebral
 malaria, activation of the HIV virus, graft-host rejection, bone
 resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
 They can also be used to treat e.g. autoimmune diseases such as multiple
 sclerosis and insulin-dependent diabetes and inflammatory and infectious
 diseases such as silicosis, and sarcoidosis, idiopathic pulmonary
 fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxic shock,
 atherosclerosis, histamine-mediated allergic reactions and immunological
 disorders including late phase allergic reactions, chronic urticaria, and
 atopic dermatitis by inhibiting chemokine-induced mast cell and basophil
 degranulation and release of histamine. IGF-mediated allergic reactions
 such as allergic asthma, rhinitis and eczema, inflammatory and
 diseases, rheumatoid arthritis, inflammation, degenerative and
 inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome,
 subepithelial basement membrane fibrosis or adult respiratory distress
 syndrome. The products can also be used for detection, diagnosis and
 drug screening.
 Sequence 1100 BP: 343 A; 259 C; 250 G; 248 T


```

Db 869 ATGAGATGTCACATTTTGGTGCAATGAAACGCTGTGACCTACTACACCAATGCTG 928
Qy 1203 tagtatttctcccttctctgtacctctctaaagaagaatcttaactgaataac 1260
Db 929 TAGCTATTTTCCCTTCTCTGTACCTCTAAGAAGAAAGAAATCTAACTGAAATAC 986

RESULT 6
V30935 ID V30935 standard; DNA; 338 BP.
AC V30935;
DT 11-SEP-1998 (first entry)
DE Homo sapiens neutrokin alpha gene related clone HSOAD55R.
KW neutrokin alpha; cell proliferation; differentiation; migration;
KW cytotoxicity; cell death; treatment; tumour; infection; inflammation;
KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
KW fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
KW cachexia; detection; diagnosis; drug screening; HSOAD55R; ss.
OS Homo sapiens.
PN WO9818921-A1.
PD 07-MAY-1998.
PF 25-OCT-1996; U17957.
PR 25-OCT-1996; WO-U17957.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ebner R, Ni J, Yu G;
DR WPI; 98-272216/24.
PT New isolated human Neutrokin alpha - used to develop products for
PT diagnosis and treatment of e.g. tumours, infections,
PT immunodeficiencies or autoimmune diseases
PT Claim 21; Page 85-86; 104pp; English.
PS The sequence is that of the neutrokin alpha related clone HSOAD55R.
CC Neutrokin alpha (NA) polypeptides modulate cell proliferation,
CC differentiation, migration, cytotoxicity and cell death.
CC They can be used to treat e.g. tumour and tumour metastasis, infections
CC by bacteria, viruses and other parasites, immunodeficiencies,
CC inflammatory diseases, lymphadenopathy, autoimmune diseases, graft
CC versus host disease and to stimulate peripheral tolerance, destroy some
CC transformed cell lines, mediate cell activation and proliferation, and
CC are functionally linked as primary mediators of immune regulation and
CC inflammatory responses. Such activity is useful for immune enhancement
CC or suppression, myeloprotection, stem cell mobilisation, acute and
CC chronic inflammatory control and treatment of leukaemia. They can also
CC be used to stimulate wound healing and to treat fibrotic disorders
CC including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They
CC can also be used to regulate haematopoiesis, by regulating the activation
CC and differentiation of various haematopoietic progenitor cells, e.g. to
CC release mature leukocytes from the bone marrow following chemotherapy, NA
CC and in stem cell mobilisation. NA may also be used to treat sepsis. NA
CC antagonists can be used to prevent septic shock, inflammation, cerebral
CC malaria, activation of the HIV virus, graft-host rejection, bone
CC resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
CC They can also be used to treat e.g. autoimmune diseases such as multiple
CC sclerosis and insulin-dependent diabetes and inflammatory and infectious
CC diseases such as silicosis, and sarcoidosis, idiopathic pulmonary
CC fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxic shock,
CC atherosclerosis, histamine-mediated allergic reactions and immunological
CC disorders including late phase allergic reactions, chronic urticaria, and
CC atopic dermatitis by inhibiting chemokine-induced mast cell and basophil
CC degranulation and release of histamine. IgE-mediated allergic reactions
CC such as allergic asthma, rhinitis and eczema, inflammatory pulmonary
CC diseases, rheumatoid arthritis, inflammation, degenerative and
CC inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome,
CC subepithelial basement membrane fibrosis or adult respiratory distress
CC syndrome. The products can also be used for detection, diagnosis and
CC drug screening.
SQ Sequence 338 BP; 90 A; 81 C; 77 G; 77 T;

Query Match 23.6%; Score 302.8; DB 1; Length 338;
Best Local Similarity 94.1%; Pred. No. 6.9e-69;
Matches 318; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

```

```

Qy 186 aggataaactctctgaggggtgagccaaagccctgcacatgtagtcagcaggacatacaac 245
Db 1 AGGNTAACTCTCTCTAGGGGTGAGCCAAAGCCCTGCCCATGTAGTCAGCAGGACATCANC 60
Qy 246 aacacagataaacaggaatgatccattccctctggtgcacttattctaaaggcccaacc 305
Db 61 AACACANNNNNCAGGAATAATCCATTCCTCTGGTGCACCTTATTCCTAAAGGCCCAACC 120
Qy 306 ttcaaatgttcaagttagtgatgatgagtcacccacagaaaggagcagtcacgccttactt 365
Db 121 TTCAAACTTCAAGTAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 366 ctgctctaagaaaaagagaagaatgaaactg-aagagtggtgtttccatctctccaccg 424
Db 181 CTTGCTCTTAAAGAAAAGAGAAATGAACCTGNAAGGAGTGTGTTTCCATCTCTCCACGG 240
Qy 425 aaggaagccctctgtccatctccacagcaggaagcgtggtgcacacttgcg 484
Db 241 AAGGAAAGCCCCCTCTNTCCGATCTCCAAAGACGGAAGCTGCTGGCTGCAACCTTGNTG 300
Qy 485 ctggcactgctgtcttctgtcctcaccggtggtgtcttt 522
Db 301 NTGGCATTTGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 338

RESULT 7
V39987 ID V39987 standard; CDNA; 617 BP.
AC V39987;
DT 03-OCT-1998 (first entry)
DE Mouse T cell surface antigen 63954 nucleic acid sequence.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;
KW diagnosis; antigen-specific proliferation; cytokine production;
KW immune response; autoimmune disorder; rheumatoid arthritis;
KW systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis; ss.
OS Mus sp.
PN WPI; 98-362719/31.
DR P-PSDB; W62463.
PT New isolated polypeptide, 63954 - used to develop products for
PT treating e.g. autoimmune disorders, inflammation, tissue rejection,
PT cancer or degenerative conditions
PT Claim 6; Page 65; 69pp; English.
PS The present sequence encodes a mouse T cell surface antigen, designated
CC 63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
CC production on effector cells and may potentiate immune cell expansion or
CC apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
CC molecule for regulation of T cell mediated cell activation, and may cause
CC a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
CC 63954 can be used to modulate immune responses in abnormal situations,
CC e.g. autoimmune disorders, including rheumatoid arthritis, systemic
CC lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
CC as acute and chronic inflammatory responses in which T cell activation,
CC expansion, and/or immunological T cell memory play an important role,
CC such as chronic inflammation or tissue rejection. The products can also
CC be used in the treatment of conditions associated with abnormal
CC physiology or development, including abnormal proliferation, e.g.
CC cancerous conditions, or degenerative conditions. The products can also
CC be used for detection, diagnosis and drug screening.
SQ Sequence 617 BP; 176 A; 156 C; 143 G; 142 T;

```


PN W09818921-A1.
PD 07-MAY-1998.
PF 25-OCT-1996; U17957.
PR 25-OCT-1996; W0-U17957.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Ebner R, Ni J, Yu G;
DR WPI: 98-272216/24.
PT New isolated human Neurokinine alpha - used to develop products for
PT diagnosis and treatment of e.g. tumours, infections,
PT immunodeficiencies or autoimmune diseases
PS Claim 21; Page 86-87; 104pp; English.
CC The sequence is that of the neurokinine alpha related clone HLTBM08R.
CC Neurokinine alpha (NA) polypeptides modulate cell proliferation,
CC differentiation, migration, cytotoxicity and cell death.
CC They can be used to treat e.g. tumour and tumour metastasis, infections
CC by bacteria, viruses and other parasites, immunodeficiencies,
CC inflammatory diseases, lymphadenopathy, autoimmune diseases, graft
CC versus host disease and to stimulate peripheral tolerance, destroy some
CC transformed cell lines, mediate cell activation and proliferation, and
CC are functionally linked as primary mediators of immune regulation and
CC inflammatory responses. Such activity is useful for immune enhancement
CC or suppression, myeloprotection, stem cell mobilisation, acute and
CC chronic inflammatory control and treatment of leukaemia. They can also
CC be used to stimulate wound healing and to treat fibrotic disorders
CC including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They
CC can also be used to regulate haematopoiesis, by regulating the activation
CC and differentiation of various haematopoietic progenitor cells, e.g. to
CC release mature leukocytes from the bone marrow following chemotherapy,
CC and in stem cell mobilisation. NA may also be used to treat sepsis. NA
CC antagonists can be used to prevent septic shock, inflammation, cerebral
CC malaria, activation of the HIV virus, graft-host rejection, bone
CC resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
CC They can also be used to treat e.g. autoimmune diseases such as multiple
CC sclerosis and insulin-dependent diabetes and inflammatory and infectious
CC diseases such as silicosis, and sarcoidosis, idiopathic pulmonary
CC fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxic shock,
CC atherosclerosis, histamine-mediated allergic reactions and immunological
CC disorders including late phase allergic reactions, chronic urticaria, and
CC atopic dermatitis by inhibiting chemokine-induced mast cell and basophil
CC degranulation and release of histamine. IGE-mediated allergic reactions
CC such as allergic asthma, rhinitis and eczema, inflammatory pulmonary
CC diseases, rheumatoid arthritis, inflammation, degenerative and
CC inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome,
CC subepithelial basement membrane fibrosis or adult respiratory distress
CC syndrome. The products can also be used for detection, diagnosis and
CC drug screening.
SQ Sequence 497 BP; 138 A; 104 C; 119 G; 119 T;

Query Match 14.8%; Score 189.4; DB 1; Length 497;
Best Local Similarity 99.0%; Pred. No. 1.1e-39;
Matches 190; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 619 caagcggcgtgaggaagctcagctgctcaccgaggactgaataatcttgaaccacc 678
|||||
DB 14 CRAAGCGCGCTGGAGGAGCTCAGCTGTCACCGGGGAGCTGAAATCTTTGAACACC 73
|||||
QY 679 agctccagagagaggcaactccagtcagacagcagataagcgtgcggttcagggtcc 738
|||||
DB 74 AGCTCCAGGAGGAGGCAACTCCAGTCAGAACAGCAGCAAAATAAGCGTTCAGGGTCC 133
|||||
QY 739 agaagaacagtcactcaagactgctgtaactgattgagacagtgaaacaccactat 798
|||||
DB 134 AGAAGAACAGTCACCTCAAGACTGCTTGAACGTGTTGCAGACAGTGAACACCACTAT 193
|||||
QY 799 acaaaaggatc 810
|||||
DB 194 ACAAAAGGCTC 205
|||||

RESULT 10
V58754
ID V58754 standard; cDNA: 1328 BP.

V58754;
18-JAN-1999 (first entry)
Human secreted protein ax318_3 cDNA.
Secreted protein; human; ax318_3; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 521..1114
FT sig_peptide /*tag= a
FT mat_peptide /*tag= b
FT 602..1111
FT /*tag= c
PN W09840404-A2.
PD 17-SEP-1998.
PF 09-MAR-1998; U04601.
PR 06-MAR-1998; US-036321.
PR 11-MAR-1997; US-815381.
PA (GENY) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 98-520802/44.
DR P-PSDB: W69423.
PT New isolated polynucleotides and secreted proteins - useful as, e.g.
PT nutritional additives, immunostimulators, haematopoiesis regulators
PT and as diagnostic agents
PS Claim 1; Page 66; 119pp; English.
CC This full-length cDNA clone, designated ax318_3, codes for a novel
CC human secreted protein (see W69423), i.e. ax318_3 protein. It was
CC isolated from a human adult testis cDNA library using methods which
CC are selective for cDNAs encoding secreted proteins, or was
CC identified as encoding a secreted or transmembrane protein on the
CC basis of computer analysis of the encoding protein. Homology is
CC shown to some database sequences. The invention provides isolated
CC polynucleotides (see V58754-63) obtained from human adult testis,
CC brain, retina or placenta, or from foetal kidney or brain cDNA
CC libraries. These are all deposited as ATCC 98553. They encode
CC novel human secreted proteins (see W69423-33) that may have e.g.
CC nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombolytic activity, receptor/ligand activity,
CC antiinflammatory activity, cadherin/tumour invasion suppressor,
CC activity, tumour inhibition or other activities. They may also be
CC used for diagnostic purposes.
SQ Sequence 1328 BP; 446 A; 307 C; 282 G; 293 T;

Query Match 3.6%; Score 46.8; DB 1; Length 1328;
Best Local Similarity 67.3%; Pred. No. 0.0089;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1187 acttacacacatgctgtagctatttccctctctctctgtaccttaagaagaagaat 1246
|||||
DB 396 AATTGTCTCCAAATGTGTAGACCATATCTCCCTGCCCTTCCCTCCCAAAAAA 455
|||||
QY 1247 ctaactgaaaaatacaaaaaaataaaaaa 1284
|||||
DB 456 AAAAAA 493

RESULT 11
V05164
ID V05164 standard; cDNA: 2705 BP.
AC V05164;
DT 06-JUL-1998 (first entry)
DE Human growth arrest gene B4B.
KW Growth arrest gene; B4B gene; cell proliferation; cancer; leukaemia;
KW marker; gene therapy; human; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 123..596
FT /*tag= a
FT /*note= "(Claim 4)"

REFERENCE 1 (bases 1 to 233)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Ruitman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
 TITLE WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilton RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 713 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 211.
 Location/Qualifiers
 1..233
 /organism="Homo sapiens"
 /note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTCAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'], [5' double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
 /db_xref="taxon:9606"
 /clone="366305"
 /clone_lib="Soares fetal heart NBHL19W"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 <1..>233
 BASE COUNT 40 a 96 c 57 g 35 t 5 others
 ORIGIN
 Query Match 39.9%; Score 183; DB 15; Length 233;
 Best Local Similarity 96.2%; Pred. No. 1.1e-36;
 Matches 227; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
 QY 136 gcatcgtgtccactgtgctggcggtgattgcccccggcaccacagacacgag 195
 Db 1 GCATCGTGTCCACCTGGTGGCGGGGTATTGCCCC-GGNACCCCGAGCAGACACGCGAG 59
 QY 196 tgccagcgtgccccccagacc-ttctcagcagagagctccagctcagagtgcca 254
 Db 60 -GCCAGCGTGGCCCCCAGGACCTTTCTCAGCAGCAGCTCCAGCTCAGAGCTGCCA 118
 QY 255 gccccacgcgaactgcagggcgctggcgctccctcaatgtgcaggctcttctccca 314
 Db 119 GCCCCCGGAACATGCACGNNCCI-GGNNYGGCCCTCAATGTGCAGGCTCTTCTCCCA 177
 QY 315 tgacacctgtgaccagctgactggttccccctcagcaccaggggtaccagag 370
 Db 178 TGACACCGTGTGCACAGCTGCTACTGGCTTCCCTCCTCAGCACCGAGGTACCANGAG 233
 RESULT 4
 AA325843 191 bp mRNA EST 20-APR-1997
 LOCUS EST28933 Cerebellum II Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION
 AA325843

NID
 KEYWORDS
 SOURCE
 ORGANISM

g1978108
 EST.
 human.
 Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominiidae;
 Homo.

REFERENCE
 AUTHORS

1 (bases 1 to 191)
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitchugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 JOURNAL
 MEDLINE
 COMMENT
 Other ESTs: THC162279
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..191
 /organism="Homo sapiens"
 /note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
 /db_xref="ATCC (inhost):126377"
 /db_xref="taxon:9606"
 /clone_lib="Cerebellum II"
 /tissue_type="cerebellum"
 /dev_stage="adult"
 <1..>191

FEATURES
 SOURCE

BASE COUNT 33 a 66 c 49 g 39 t 4 others
 ORIGIN

mRNA
 BASE COUNT
 ORIGIN

Query Match 39.0%; Score 179; DB 17; Length 191;
 Best Local Similarity 95.3%; Pred. No. 1.1e-35;
 Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 262 cgaactgcacggccctggcgccctcaatgtgcaggctcttctccatgcacac 321
 Db 1 CGCAACTGCACGGCCCTGGGAGCTGGCCCTCAATGTGCAGGNTCTTCTCCCATGACAC 60
 QY 322 ctgtgcaccagctgcactgcttccccctcagcaccaggtaccaggtgagagtgt 381
 Db 61 CTGTGCACCACTGCACCTGGCTTCCCCCTCAGCACCGGTACCANGAGCTTGAGGAGTGT 120
 QY 382 gacgctgccgtcatcagcttggcttccagacatctccatcaagaggtcagcgg 441
 Db 121 GAGCNTCCCTCATCGACTTTTGGCTTTCCAGGACATCTCCATCAAGAGGCTGCACGG 180
 QY 442 ctgctgcaggc 452

```

Db 181 CTGCTCANGCC 191
|||||
RESULT 5
AA025673/c 467 bp mRNA EST 01-FEB-1997
LOCUS ze90h09.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
DEFINITION 36305 3', mRNA sequence.
ACCESSION AA025673
NID g1491066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Eukaryotes; Chordata; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 467)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterston,R., Williamson,A., Wohlmann,P. and Willson,R.
TITLE WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 713 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 373.
FEATURES
Location/Qualifiers
1..467
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCATCTTTTCTTTTCTTTT 3'], and
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHH19W."
/db_xref="taxon:9606"
/clone="366305"
/clone_lib="Soares fetal heart NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>467)
BASE COUNT 94 a 135 c 134 g 99 t 5 others
ORIGIN
mRNA
Query Match 32.5%; Score 149; DB 15; Length 467;
Best Local Similarity 95.6%; Pred. No. 3.4e-28;
Matches 152; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 300 aggccttcctccatgacacacctgtgacacagctgacagcttcctccctccagcaccag 359
Db 467 AGNCCTCTTCTTCTCCCATGACACCTGTGACACAGCTGCATGTGGCTTCCCTCCAGCACCAG 408

Qy 360 ggtaccagagctgaggagctgagcgtgccctcatcacttctggtcttccagacat 419
Db 407 GNTACCAGAGCTGAGGAGTGTGACGGTCCCTCATCGACTTTGTGGCTTCCAGGACAT 348

```

```

Qy 420 ctccatcaagagctgacggcgtgctgcagggccctcga 458
|||||
Db 347 CTCCATCAAGAGGCTCAGCGGCTGCTGCAAGCCCTCGA 309

RESULT 6
AI290210/c 436 bp mRNA EST 30-NOV-1998
LOCUS q179g12.x1 Soares_NhMPu_S1 Homo sapiens cDNA clone IMAGE:1878598
DEFINITION 3', mRNA sequence.
ACCESSION AI290210
NID g3931864
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 436)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 388.
FEATURES
Location/Qualifiers
1..436
/organism="Homo sapiens"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not 1;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
/db_xref="taxon:9606"
/clone="IMAGE:1878598"
/clone_lib="Soares_NhMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
BASE COUNT 97 a 136 c 124 g 79 t
ORIGIN

```

```

Query Match 25.9%; Score 119; DB 23; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 gcttcctccctcagcaccagggtaccagagctgaggagctgagcgtccctcactcagact 400
Db 436 GCTTCCCTCCCTCAGCACCAGGTTACCCAGGAGTGTGAGCGTCCGCTCATCGACT 377

Qy 401 ttgtggcttccagagacatcccatcaagagcgtgcagcgtctgcagccctcag 459
Db 376 TTGTGGCTTTCAGGACATCTCCATCAAGAGGCTGCACGGCTGCTGCAGGCCCTCGAG 318

RESULT 7
AA740147/c 461 bp mRNA EST 08-FEB-1998
LOCUS ob26a08.s1 NCI-CCAP_Kid5 Homo sapiens cDNA clone IMAGE:1324790,
DEFINITION mRNA sequence.

```

```

ACCESSION   AA740147
NID         92778739
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            unknown library type
            Insert Length: 1402 Std Error: 0.00
            Seq primer: -40ml3 fwd. ET from Amersham.
FEATURES   source
            Location/Qualifiers
            1..461
            /organism="Homo sapiens"
            /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
            a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5'-
            AACTGGAAGAATTGCGCGCGCAATATTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M. Fatima Bonaldo."
            /db_xref="taxon:9606"
            /clone_lib="NCI-CGAP_Kid5"
            /tissue_type="2 pooled tumors (clear cell type)"
            /lab_host="DH10B"
            84 a 148 c 158 g 71 t
            BASE COUNT
            ORIGIN

Query Match 18.0%; Score 82.8; DB 26; Length 461;
Best Local Similarity 92.6%; Pred. No. 1.2e-11;
Matches 87; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 365 cagagctgagagtgtagctgctgcgcctgcgcctcga 424
|||||
Db 371 CAGGAGCTGAGAGTGTCGAGGTCGCCGTCATCGACTTTGTGCTTCCAGGACATCTCCA 312
|||||

Qy 425 tcaagagctgcagcggctgctgcaggccctcga 458
|||||
Db 311 TCAAGAGCGCTGAGCGGCTGCTGCCAGCCCTCGA 278
|||||

RESULT 8
LOCUS      A1185297/c
DEFINITION qe36c02.s1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:1741038 3', mRNA sequence.
ACCESSION  A1185297
NID        93735935
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 342.
Location/Qualifiers
1..344
/organism="Homo sapiens"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot - 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library. Soares
fetal heart NbHL19W."
/db_xref="taxon:9606"
/clone="IMAGE:1741038"
/clone_lib="Soares_fetal_lung_NbHL19W"
/lab_host="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
72 a 115 c 106 g 51 t
BASE COUNT
ORIGIN

Query Match 17.0%; Score 78; DB 30; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 382 gagcgtgcgcctgcgcctgctgttcgcagacatctccatcagagcgtgcagcgg 441
|||||
Db 344 GAGCGTGCCTGCATCGACTTTGTGCTTCCAGGACATCTCCATCAGAGGCTGCAGCGG 285
|||||

Qy 442 ctgctgcagcgcctgcag 459
|||||
Db 284 CTGCTGCAGGCCCTCGAG 267
|||||

RESULT 9
LOCUS      AA613366/c
DEFINITION ng31a08.s1 NCI-CGAP_Col0 Homo sapiens cDNA clone IMAGE:1145462,
            mRNA sequence.
ACCESSION  AA613366
NID        92464404
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 978 Std Error: 0.00

```

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 367.

FEATURES

source

Location/Qualifiers
1..378
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Benito Soares and M. Fatima Bonaldo (N-Soares4)."
/db_xref="taxon:9606"
/clone="IMAGE:1145462"
/clone_lib="NCI_CGAP_Co10"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
BASE COUNT 90 a 117 c 109 g 62 t
ORIGIN

Query Match 14.6%; Score 67.2; DB 24; Length 378;
Best Local Similarity 95.8%; Pred. No. 9.7e-08;
Matches 69; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 388 gcgcgtacgaatttgcttcaggacatccatcaaggaggtgcggctgctg 447
|||||
Db 378 GCGGTCATCGACTTTGTGGCTTTCAGGACATCTCCATCAAGAGGCTCGAGGGCTGCTG 319
|||||

QY 448 caggccctcgag 459
|||||

Db 318 CATGCCCTCGAG 307
|||||

RESULT 10
A1282851/C
LOCUS
DEFINITION qt84e09.x1 NCI_CGAP_Co14 Homo sapiens cDNA clone IMAGE:1961992 3',
mRNA sequence.
ACCESSION A1282851
NID g3921084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 323.

FEATURES

source

Location/Qualifiers
1..358
/organism="Homo sapiens"
/note="Organ: colon; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.7 kb. Life Technologies catalog #:

11531-019"
/db_xref="taxon:9606"
/clone="IMAGE:1961992"
/clone_lib="NCI_CGAP_Co14"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 81 a 114 c 103 g 60 t
ORIGIN

Query Match 10.5%; Score 48; DB 23; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 caggacatccatcaagaggtgcggctgctgcaggccctcgag 459
|||||
Db 358 CAGGACATCTCCATCAAGAGGCTCGAGGGCTGCTGCAGGCCCTCGAG 311
|||||

RESULT 11

AA210969

LOCUS

DEFINITION AA210969 452 bp mRNA EST 31-JAN-1997
zq88c08.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
649070 5' similar to gb:J03827 Y BOX BINDING PROTEIN-1 (HUMAN);,
mRNA sequence.

ACCESSION

NID

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 452)

AUTHORS

Hillier, L., Clark, N., Duboue, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
WashU-Merck EST Project
Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 331.
Location/Qualifiers
1..452

FEATURES

source

/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Cloned unidirectionally. Primer: Oligo dt.
Differentiated, post mitotic hNT neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'
CTCCAGCTTTTCTTTTCTTTT 3'."
/db_xref="taxon:9606"
/clone="649070"
/clone_lib="Stratagene hNT neuron (#937233)"
/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 113 a 142 c 128 g 59 t 10 others
ORIGIN

mrna

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 9.2%; Score 42.2; DB 15; Length 452;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

This Page Blank (uspto)

ALIGNMENTS

QY	1	ctgagcgctgcgctactgcaacgctcctctgcggggagcgtgagaggaagcagcgagct	60
Db	277	CTGAGCGCTGCCGCTACTGCAAGCTCTTCGGGGAGCGTGAGGAGGACGGCT	336
QY	61	tgcacgccaccacaacgctgctgcgctgcgcacccgctctctgcgcagctagt	120
Db	337	TGCACGCCACCCACAACCGTCTGCGCTGCGCACCGGCTTCTTCGCGACGCTGGT	396
QY	121	tctgcttgagcagcgcgcgtgtccacctggtgccgcgctgattgcccggggcaccccc	180
Db	397	TTCTGCTTGAGCAGCGCATCGTGTCCACCTTGTCGGCGCTGATTGCCCGGGACCCCC	456
QY	181	agccagaacacgcagtgccagcgtgccccccagggacctctcagccagcagctccagc	240
Db	457	AGCCAGAACACGAGTGCAGCGCTGCCCGCAGGCACCTTCTCAGCCAGCAGCTCCAGC	516

QY 10 tgcgcctactgcaacgctctctgcccggagcgtgaggagagcagcggttgccacgcc 69
 DB 361 TCGGTGTACTGACGCCCGTGTGCAGGAACCTGCAGACCTGGAACAGAGTGCAACGCC 420
 QY 70 acccacaacccgtgcgcgtgcgcgcacccgctctcttctgcgcacgctggttctgcttg 129
 DB 421 ACCCACAACCGAGTGTGCGAATGTGAGGAAGGCGCTACTCTGGAGCTCGAATCTGCTTG 480
 QY 130 gagcagcgcgtgctccacactggtgcccggcgtgattgccccgggacccccagccagAAC 189
 DB 481 AAGCACCAGGAGTGTCCCCAGCGCTGGGTGTGCTGCAGGCTGGACCCAGAGGGAAC 540
 QY 190 acgagcgtgcagcgtgccccccagcagcacttctcagcagcagctccagctcagagcag 249
 DB 541 ACGTTTGCAGAAAGATGCCGATGGTCTTCTCAGGTGAGACGTATCGAAAGCACCC 600
 QY 250 tgcagcgcacccagcaactgcagcgcgtggcctggcctcaatgtgcagcgtctctcc 309
 DB 601 TGTAGGAACACACCAACTGCAGCTCACTTGGCCCTCTGCTAATTCAGAAAGGAATGCA 660
 QY 310 tcccatgacaccctgtgcaccagctgca 337
 DB 661 ACACATGCAATGTATGTTCGGAACA 688
 RESULT 6
 V39094 ID V39094 standard; cDNA; 497 BP.
 AC V39094;
 DT 26-OCT-1998 (first entry)
 DE HELDIO6R DNA sequence.
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue; HELDIO6R DNA sequence; ss.
 OS Homo sapiens.
 PN WO9830694-A2.
 PD 16-JUL-1998.
 PF 13-JAN-1998; U00153.
 PR 14-JAN-1997; US-035496.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 DR WPI: 98-399142/34.
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 PS Disclosure: Fig 6; 91pp; English.
 CC The present sequence represents the HELDIO6R DNA sequence which is
 CC related to extensive portions of the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha; V39085) cDNA and the TNFR-6 beta cDNA
 CC (V39086). The invention provides for TNFR-6 alpha cDNA and TNFR-6 beta
 CC cDNA, and the TNFR-6 alpha (W63622) and TNFR-6 beta (W63623) proteins
 CC they encode, respectively. TNFR-6 alpha and TNFR-6 beta are members of
 CC the tumour necrosis factor receptor (TNFR) family. TNFRs are expressed
 CC in endothelial cells, keratinocytes, normal prostate and prostate tumour
 CC tissue. For a number of disorders of these cells, particularly of the
 CC immune system, substantially altered (whether increased or decreased)
 CC levels of TNFR-6 alpha and/or TNFR-6 beta gene expression can be
 CC detected, therefore the TNFR-6 alpha and TNFR-6 beta polypeptides,
 CC nucleic acids and antibodies are claimed to be useful in the diagnosis
 CC of such disorders. Mutations of the TNFR-6 alpha and TNFR-6 beta genes
 CC can also be detected. The TNFR polypeptides are also claimed to be
 CC useful for identifying ligands which may be useful in the treatment of
 CC apoptosis related disorders.
 SQ Sequence 497 BP; 79 A; 143 C; 150 G; 92 T;
 Query Match 16.9%; Score 77.8; DB 1; Length 497;
 Best Local Similarity 76.7%; Pred. No. 2e-09; Mismatches 29; Indels 1; Gaps 1;
 Matches 99; Conservative 0;
 QY 1 ctggagcgtgcgcgtactgcaactctctctgcccggagcgtgaggagagcagcggttg 60
 DB 309 CTGGAGCGCTGNCCTTACTNCAACGCTCTCTCTCGGGAGCGTTCAGGAGGACGCGGTT 368

QY 61 tgcacgcccacccacacacgt-gcctgcgcgtgcgcgcacccggtcttctgcgcagcgtgg 119
 DB 369 TNCCACGCAACCAACACCGGNTTACCCTNGCGNACCGGTTTCTTCGNGGCAAGTTG 428
 QY 120 ttcttgctt 128
 DB 429 GTTTTNT 437
 RESULT 7
 T96062 ID T96062 standard; cDNA; 1324 BP.
 AC T96062;
 DT 20-APR-1998 (first entry)
 DE Mouse osteoprotegerin cDNA.
 KW Osteoprotegerin; antibody; diagnosis; affinity purification;
 KW recombinant production; transgenic animal; treatment; prevention;
 KW antisense oligonucleotide; probe; detection; screening; mouse;
 KW bone disease; osteoporosis; Paget's disease; hypercalcaemia;
 KW hyperparathyroidism; rheumatoid arthritis; osteomyelitis;
 KW osteolytic metastasis; periodontal bone loss; bone necrosis;
 KW osteopaenia; murine; ss.
 OS Mus sp.
 FH Key.
 FT Key Location/Qualifiers
 FT CDS 90..1295
 FT /tag= a
 FT /product= osteoprotegerin
 PN DE19654610-AL.
 PD 26-JUN-1997.
 PF 20-DEC-1996; 054610.
 PR 03-SEP-1996; US-706945.
 PR 22-DEC-1995; US-577788.
 PA (AMGE-) AMGEN INC.
 PI Boyle WJ, Calzone FJ, Lacey DL, Chang MS;
 DR WPI: 97-334271/31.
 DR P-PSDB; W38344.
 PT Nucleic acid encoding osteoprotegerin - useful for treatment of
 PT diseases involving excessive bone loss, e.g. osteoporosis
 PS Claim 1; Pages 106-107; 182pp; German.
 CC The present sequence encodes mouse osteoprotegerin (OPG). Anti-OPG
 CC antibodies can be used in OPG diagnostic assays, and as affinity
 CC purification materials. The OPG cDNA can be used to express
 CC recombinant OPG and to generate transgenic animals. It can also
 CC be used to regulate the level of OPG in mammals, specifically to
 CC increase OPG levels, however the use of antisense sequences is
 CC also contemplated. Fragments of the cDNA can be used as probes to
 CC detect OPG expressing cells and tissue, and to screen cDNA
 CC libraries for related sequences. OPG can be used to treat or
 CC prevent bone diseases, specifically excessive bone loss, e.g.
 CC osteoporosis, Paget's disease, hypercalcaemia,
 CC hyperparathyroidism, rheumatoid arthritis, Osteomyelitis,
 CC osteolytic metastases, periodontal bone loss, bone necrosis and
 CC osteopaenia.
 SQ Sequence 1324 BP; 381 A; 329 C; 335 G; 279 T;
 Query Match 16.5%; Score 75.6; DB 1; Length 1324;
 Best Local Similarity 52.2%; Pred. No. 6.8e-09;
 Matches 168; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
 QY 16 tactgcaacgtctctgcccggagcgtgaggagcagcggttgccacgccaccac 75
 DB 333 TATTGCAGCCCGAGTGTGCAAGGAACTGCGTCCGTGAAGCAGAGTGCACCGCCAC 392
 QY 76 aacgtgctccgctgcgcgcacccgcttcttcgcgcagcgtggttctgttgagcac 135
 DB 393 AACCGAGTGTGTGAGTGTGAGAAAGGCGGTGTACCTGGAGATCGAATCTGTGAAGCAC 452
 QY 136 gcattgtctccacgtggtgcccggcgtgattgccccgggcacccccagcagacacgcag 195
 DB 453 CGGAGCTGTCCCGGGGTCTCCGGCGTGTGCAAGTGTGCAACCCCGAGAGCGGAACACAGTT 512
 QY 196 tgcacgctgccccccccagccacctcttcacccagcagctccagcagcagtcagtcag 255


```

Db 513 TGCAAAAATGTCAGATGGTTCTTCTCAGGTGAGACTTCAGAAAGCACCTGTATA 572
QY 256 cccacgcgaactgcagcgccctggccctgacctcaatgtgccaggctcttcccat 315
Db 573 AAMACAGCAACTGCAGCACATTGGCCCTCCTGCTAATTCAGAAAGGAATGCAACACAT 632
QY 316 gacacctgtgcaccagctgca 337
Db 633 GACAAGTGTGTTCCGGARACA 654

RESULT 8
T33180
ID T33180 standard; DNA; 564 BP.
AC T33180;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..561 /*tag= b
FT /*product= OCIF-CSph
PN WO9626217-Al.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Iano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99949.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 84; Page 149-150; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-CSph in which amino acids 167-380 of the mature
CC protein have been deleted. These amino acid changes have been caused
CC by the introduction of a restriction site. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 564 BP; 159 A; 149 C; 132 G; 124 T;

```

```

Query Match 15.0%; Score 69; DB 1; Length 564;
Best Local Similarity 50.8%; Pred. No. 1.8e-07;
Matches 165; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaacgtctctgctggggagcggtgagaggagcgagcggttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGCTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTCAATCGC 297
QY 70 acccaacaacgtgcctgcgcgtgcgcacgcggcttcttcgcgcacgtggttctcttg 129
Db 298 ACCCAACAACGGCGTGTGCGAATGCAAGGAGCGCGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gaggcagcatggttcacactggtgcgcgcgtgattgccccgggacccccagccagaac 189
Db 358 AAACATAGAGAGTGCCTCCTTGAGTTGGAGTGGTGCAGAGTGGAAACCCAGACGGAAT 417
QY 190 acgcagtgccagcgtgccccccagccaccttcttcagccagcagctccagctcagagcag 249

```

```

Db 418 ACAGTTTGCAAAAGATGTCAGATGGTTCTTCTCAATGAGACGTCATCTAAAGCACCC 477
QY 250 tgccagcccccacgcgaactgcagcgccctggccctgacctcaatgtgccaggctcttec 309
Db 478 TGTAGAAAACACACAATTCAGTGTCTTTGGTCTCCTGCTAACTCAGAAAGGAATGCA 537
QY 310 tcccatgacacccctgtgcaccagct 334
Db 538 ACACAGCAACAATATGTTCCGGCT 562

RESULT 9
T33179
ID T33179 standard; DNA; 966 BP.
AC T33179;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..963 /*tag= b
FT /*product= OCIF-CSph
PN WO9626217-Al.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Iano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99949.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 81; Page 149; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-CSph in which amino acids 298-380 of the mature
CC protein have been deleted and replaced by Ser-Leu-Asp. These amino
CC acid changes have been caused by the introduction of a restriction
CC site. The OCIF of the invention has a molecular weight by SDS-PAGE of
CC 60 kD under reducing conditions and 120 kD under non-reducing
CC conditions. The protein is adsorbed onto cation-exchangers or heparin
CC and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56
CC deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful in the
CC control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 966 BP; 301 A; 228 C; 226 G; 211 T;

```

```

Query Match 14.9%; Score 68.6; DB 1; Length 966;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaacgtctctgctggggagcggtgagaggagcgagcggttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGCTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTCAATCGC 297
QY 70 acccaacaacgtgcctgcgcgtgcgcacgcggcttcttcgcgcacgtggttctcttg 129
Db 298 ACCCAACAACGGCGTGTGCGAATGCAAGGAGCGCGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gaggcagcatggttcacactggtgcgcgcgtgattgccccgggacccccagccagaac 189
Db 358 AAACATAGAGAGTGCCTCCTTGAGTTGGAGTGGTGCAGAGTGGAAACCCAGACGGAAT 417
QY 190 acgcagtgccagcgtgccccccagccaccttcttcagccagcagctccagctcagagcag 249
Db 418 ACAGTTTGCAAAAGATGTCAGATGGTTCTTCTCAANTGAGACGTCATCTAAAGCACCC 477

```

QY 250 tgcagccccacagcaactgcagcgccctggcgccctcaatgtgccaggtctttcc 309
Db 478 TGTAGAAAACACACAAATGCAGTGTCTTTGGTCTCTGCTAACTCAGAAAGGAATGCA 537
QY 310 tcccatgacacctgtgcacagctgcactg 340
Db 538 ACACAGACACATATGTTCCGGAACAGTG 568

RESULT 10
T33163
ID T33163 standard; DNA; 1206 BP.
AC T33163;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C21S, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis; ss.
OS Synthetic.

FT Key Location/Qualifiers
FT signal_peptide 1..63
FT mat_peptide 64..1203
FT /tag= a
FT /tag= b
FT /product= OCIF-C21S

PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99933.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 33; Page 134-135; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-C21S in which the 21st Cys residue in the mature
CC OCIF protein is substituted by Ser. The OCIF of the invention has a
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and
CC 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 1206 BP; 389 A; 286 C; 267 G; 264 T;

Query Match 14.9%; Score 68.6; DB 1; Length 1206;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaactgtctctgcggggagcggtgagaggaggagcagggcttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTGCATCGC 297
QY 70 accacaacgtgctgcgcgtgcgcacggcttcttcgcgcagcgtggtttctgttg 129
Db 298 ACCCACAACCCGGTGTGCGAATGCAAGGAGGCGCTACCTTGAGATAGAGTCTTGCTTG 357
QY 130 gagcagcagcgtgtccacctgggtgcggcggtgattccccgggcacccccagccagaac 189
Db 358 AACATAGGAGCTGCCCTCTCTGGATTGGAGTGGTGCAGCTGGAACCCAGAGCGGAAT 417
QY 190 acgagtgccagcgctgccccccagggcaccttctcagccagcagctccagctcagagcag 249
Db 418 ACAGTTTGCAAAAGATGCCAGATGGGTCTCTCAAAATGAGAGCTCATCTAAAGCACCC 477
QY 250 tgcagccccacagcaactgcagggccctggcgccctcaatgtgccaggtctttcc 309

Db 478 TGTAGAAAACACACAAATGCAGTGTCTTTGGTCTCTGCTAACTCAGAAAGGAATGCA 537
QY 310 tcccatgacacctgtgcacagctgcactg 340
Db 538 ACACAGACACATATGTTCCGGAACAGTG 568

RESULT 11
T33164
ID T33164 standard; DNA; 1206 BP.
AC T33164;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C22S, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis; ss.
OS Synthetic.

FT Key Location/Qualifiers
FT signal_peptide 1..63
FT mat_peptide 64..1203
FT /tag= a
FT /tag= b
FT /product= OCIF-C22S

PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99934.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 36; Page 135-136; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-C22S in which the 22nd Cys residue in the mature
CC OCIF protein is substituted by Ser. The OCIF of the invention has a
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and
CC 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 1206 BP; 389 A; 285 C; 268 G; 264 T;

Query Match 14.9%; Score 68.6; DB 1; Length 1206;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaactgtctctgcggggagcggtgagaggaggagcagggcttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTGCATCGC 297
QY 70 accacaacgtgctgcgcgtgcgcacggcttcttcgcgcagcgtggtttctgttg 129
Db 298 ACCCACAACCCGGTGTGCGAATGCAAGGAGGCGCTACCTTGAGATAGAGTCTTGCTTG 357
QY 130 gagcagcagcgtgtccacctgggtgcggcggtgattccccgggcacccccagccagaac 189
Db 358 AACATAGGAGCTGCCCTCTCTGGATTGGAGTGGTGCAGCTGGAACCCAGAGCGGAAT 417
QY 190 acgagtgccagcgctgccccccagggcaccttctcagccagcagctccagctcagagcag 249
Db 418 ACAGTTTGCAAAAGATGCCAGATGGGTCTCTCAAAATGAGAGCTCATCTAAAGCACCC 477
QY 250 tgcagccccacagcaactgcagggccctggcgccctcaatgtgccaggtctttcc 309
Db 478 TGTAGAAAACACACAAATGCAGTGTCTTTGGTCTCTGCTAACTCAGAAAGGAATGCA 537

```
QY 310 tccatgacaccctgtgcaccagctgactg 340
    ||| ||| ||| ||| ||| ||| |||
Db 538 ACACAGCAACATATGTTCCGGAAACAGTG 568

RESULT 12
T33166
ID T33166 standard; DNA; 1083 BP.
AC T33166;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR1, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63
FT mat_peptide 64..1080
FT /tag= a
FT /tag= b
FT /product= OCIF-DCR1
PN W09626217-Al.
PD 29-AUG-1996. J00374.
PF 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99936.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 42; Page 137-138; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-DCR1 in which amino acids 2-42 of the mature
CC protein have been deleted. The OCIF of the invention has a molecular
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
CC in the control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 1083 BP; 352 A; 250 C; 246 G; 235 T;

Query Match 14.9%; Score 68.6; DB 1; Length 1083;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tcccgtactgcaagctcctctgcggggagcgtgaggagggcagcggttgcaccgcc 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TGTCTATATGTCAGCCCGCTGTGCAAGGAGCTGCATGCTCAAGCAGGAGTGCAATCGC 174

QY 70 acccaacaacgtgcctgcgcgtgcgcacacggcttcttcgcgcacgctggtttctgtg 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ACCCAACACCGGTGTGCCAATGCAAGGAGCGCTACTCTGAGATAGAGTTCTGCTTG 234

QY 130 gaggcagcatcgtgcacctggtgcgcgttccttcgcgcacgctggtttctgtg 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 ACCCAACACCGGTGTGCCAATGCAAGGAGCGCTACTCTGAGATAGAGTTCTGCTTG 234

QY 130 gaggcagcatcgtgcacctggtgcgcgttccttcgcgcacgctggtttctgtg 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 AAACATAGAGAGTGCCTCTCTGATTTGGATGGTGTCAGAGTGGAAACCCAGAGCGCAAT 294

QY 190 agcagtcgacccgctgccccccagcacccttcacgagcagcagcagcagcagcag 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 AAACATAGAGAGTGCCTCTCTGATTTGGATGGTGTCAGAGTGGAAACCCAGAGCGCAAT 294

QY 190 agcagtcgacccgctgccccccagcacccttcacgagcagcagcagcagcagcag 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 ACAGTTTGCAAAAGATGTCAGATGTCCTTCTGATTTGGTCTCTCTGCTAAAGCAACCC 354

QY 250 tccaccccccacacgaactgcacggccctggcctggcctgaatgtgccagctcttcc 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 ACAGTTTGCAAAAGATGTCAGATGTCCTTCTGATTTGGTCTCTCTGCTAAAGCAACCC 354

QY 250 tccaccccccacacgaactgcacggccctggcctggcctgaatgtgccagctcttcc 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TGTAGAAACACACAAATTCAGTGTCTTTGGTCTCTCTGCTAAAGCAACCCAGCAATGCA 414

QY 310 tccatgacaccctgtgcaccagctgactg 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 14
T33171
ID T33171 standard; DNA; 984 BP.
AC T33171;
DE Mutated OCIF, OCIF-DDD2, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63
FT mat_peptide 64..981
FT /tag= a
FT /tag= b
FT /product= OCIF-DDD2
WO9626217-A1.
29-AUG-1996. J00374.
20-FEB-1996; JP-054977.
21-JUL-1995; JP-207508.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB: R99941.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 57; Page 142-143; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-DDD2 in which amino acids 253-326 of the mature
CC protein have been deleted. The OCIF of the invention has a molecular
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
CC in the control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 984 BP; 313 A; 236 C; 212 G; 223 T;

Query Match 14.9%; Score 68.6; DB 1; Length 984;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaactctctctgaggagcgtgagagagcagcgggttgcacgcc 69
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
238 TGTCTATACTGTCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAGAGGAGTGCAATCGC 297
QY 70 acccaaacctgctgcgcgtccgcacgcggttcttcgcgcacgctgtttctgttg 129
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 ACCCAACCGCGTGTGCGAATGCAAGGAGGCGGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gagcagcgtctgttccacctggtgcgcggtgattgccccgggccccccagcagaac 189
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 AAACATAGGAGCTGCCCTCTCTGGATTGGAGTTGGAGTGGTGAAGCTGGAACCCAGAGCAAAAT 417
QY 190 acgcagtgccagccgtgccccccagggcaccttcttcagccagcagctccagtcagagcag 249
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
418 ACAGTTTGAACAAAGATGTCAGATGGTCTTCTCAAAATGAGAGCTCATCTAAAGCACCC 477
QY 250 tgcagccccccagcagcagccctggcctggcctccatgagcagcagcttcc 309
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
478 TGTAGAAAACACACAAATTCAGTGTCTTTGGTCTCTCTGCTAACTCAGAAGGAATGCA 537
QY 310 tcccatgacacctgtgcacagctgacctg 340
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
538 ACACAGCACACATATGTTCCGGAAACAGTG 568

```

```

RESULT 15
T33173
ID T33173 standard; DNA; 1056 BP.

```

```

AC T33173;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CC, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..83
FT mat_peptide 84..1053
FT /tag= a
FT /tag= b
FT /product= OCIF-CC
WO9626217-A1.
29-AUG-1996. J00374.
20-FEB-1996; JP-054977.
21-JUL-1995; JP-207508.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB: R99943.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 63; Page 144-145; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-CC in which amino acids 331-380 of the mature
CC protein have been deleted. The OCIF of the invention has a molecular
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
CC in the control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 1056 BP; 332 A; 252 C; 247 G; 225 T;

Query Match 14.9%; Score 68.6; DB 1; Length 1056;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaactctctctgaggagcgtgagagagcagcgggttgcacgcc 69
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
238 TGTCTATACTGTCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAGAGGAGTGCAATCGC 297
QY 70 acccaaacctgctgcgcgtccgcacgcggttcttcgcgcacgctgtttctgttg 129
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 ACCCAACCGCGTGTGCGAATGCAAGGAGGCGGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gagcagcgtctgttccacctggtgcgcggtgattgccccgggccccccagcagaac 189
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 AAACATAGGAGCTGCCCTCTCTGGATTGGAGTTGGAGTGGTGAAGCTGGAACCCAGAGCAAAAT 417
QY 190 acgcagtgccagccgtgccccccagggcaccttcttcagccagcagctccagtcagagcag 249
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
418 ACAGTTTGAACAAAGATGTCAGATGGTCTTCTCAAAATGAGAGCTCATCTAAAGCACCC 477
QY 250 tgcagccccccagcagcagccctggcctggcctccatgagcagcagcttcc 309
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
478 TGTAGAAAACACACAAATTCAGTGTCTTTGGTCTCTCTGCTAACTCAGAAGGAATGCA 537
QY 310 tcccatgacacctgtgcacagctgacctg 340
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
538 ACACAGCACACATATGTTCCGGAAACAGTG 568

```

Search completed: May 14, 1999, 11:42:04
Job time: 5233 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:30:50 ; Search time 2254.51 seconds
(without alignments)
728.360 Million cell updates/sec

Title: US-09-212-270-7

Perfect score: 459

Sequence: 1 ctggagcgctgcgcgtactg.....ggctgctgcagggccctcgag 459

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl1.*
- 9: gb_pl2.*
- 10: gb_pr1.*
- 11: gb_pr2.*
- 12: gb_pr3.*
- 13: gb_ro.*
- 14: gb_st.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_v1.*
- 18: gb_htg.*
- 19: em_ba.*
- 20: em_fun.*
- 21: em_hum1.*
- 22: em_hum2.*
- 23: em_in.*
- 24: em_on.*
- 25: em_or.*
- 26: em_ov.*
- 27: em_pat.*
- 28: em_ph.*
- 29: em_pl.*
- 30: em_ro.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_v1.*
- 34: em_htg.*
- 35: em_sts.*
- 36: gb_ba1.*
- 37: gb_ba2.*
- 38: gb_pl1.*
- 39: gb_pl2.*
- 40: gb_pr1.*
- 41: gb_pr2.*
- 42: gb_pr3.*
- 43: gb_sts.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
-----	-----	-----	-----	-----	-----	-----	-----

1	83.2	18.1	2432	13	RNU94330	U94330 Rattus norv
2	75.6	16.5	1325	13	MMU94331	U94331 Mus musculu
3	74	16.1	2818	13	AB013898	AB013898 Mus muscu
4	68.6	14.9	1206	11	AB002146	AB002146 Homo sapi
5	68.6	14.9	1356	11	HSU94332	U94332 Human osteo
6	68.6	14.9	1206	41	AB002146	AB002146 Homo sapi
7	68.6	14.9	1356	41	HSU94332	U94332 Human osteo
8	52.8	11.5	1388	13	MMTNFR2A	X76401 M.musculus
9	51.2	11.2	3795	13	MUSTNFR1	M59378 Murine tumo
10	51.2	11.2	1505	13	MUSTNFR2	M60469 Mouse tumor
11	48.6	10.6	2116	13	AF019046	AF019046 Mus muscu
12	47.2	10.3	1968	12	AF068868	AF068868 Homo sapi
13	47.2	10.3	1968	42	AF068868	AF068868 Homo sapi
14	45	9.8	2339	6	A26415	A26415 cDNA fragme
15	44.4	9.7	1600	6	A43530	A43530 Sequence 1
16	44.4	9.7	3386	11	HUMNGFR	M14764 Human nerve
17	44.4	9.7	3386	41	HUMNGFR	M14764 Human nerve
18	43.8	9.5	9900	11	AB008821S2	AB008822 Homo sapi
19	43.8	9.5	9900	41	AB008821S2	AB008822 Homo sapi
20	43.6	9.5	1080	10	HUMTNFRA	M35857 Human tumor
21	43.6	9.5	1080	40	HUMTNFRA	M35857 Human tumor
22	43.4	9.5	1641	6	I36196	I36196 Sequence 1
23	43.4	9.5	1557	6	I36197	I36197 Sequence 3
24	43.4	9.5	691	6	I36350	I36350 Sequence 12
25	43.4	9.5	7218	6	I66494	I66494 Sequence 14
26	43.4	9.5	3683	10	HUMNFR	M32315 Human tumor
27	43.4	9.5	2394	10	HUMTNFRII	M55994 Human tumor
28	43.4	9.5	3492	10	S63368	S63368 tumor necro
29	43.4	9.5	380	13	AB013899S2	AB013900 Mus muscu
30	43.4	9.5	3683	40	HUMNFR	M32315 Human tumor
31	43.4	9.5	2394	40	HUMTNFRII	M55994 Human tumor
32	43.4	9.5	3492	40	S63368	S63368 tumor necro
33	43.4	9.5	3380	43	G26865	G26865 human STS S
34	43.2	9.4	7412	17	PVUL50S	X87246 Pseudorabie
35	43	9.4	4080	17	ADVR	X72087 Aujeszky's
36	43	9.4	5657	17	PVPCA	X80797 Pseudorabie
37	42.8	9.3	4000	13	MUSPOUDOMA	M88299 Mouse brain
38	42.6	9.3	5823	3	DROZFH1	M63449 D.melanogas
39	42.6	9.3	2168	10	HUMPSDBPB	I37516 Homo sapien
40	42.6	9.3	2168	40	HUMPSDBPB	I37516 Homo sapien
41	42.4	9.2	2461	6	I23785	I23785 Sequence 1
42	42.4	9.2	2461	10	HUMGAS	L13720 Homo sapien
43	42.4	9.2	2461	40	HUMGAS	L13720 Homo sapien
44	42	9.2	3060	3	AF027735	AF027735 Nephila c
45	41.6	9.1	28826	1	MTV041	AL021958 Mycobacte

ALIGNMENTS

RESULT 1

RNU94330	RNU94330	2432 bp	mRNA	ROD	06-MAY-1997
LOCUS	Rattus norvegicus osteoprotegerin (OPG) mRNA, complete cds.				
DEFINITION					
ACCESSION	U94330				
NID	92072180				
KEYWORDS					
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
AUTHORS	Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;				
	Murinae; Rattus.				
	1 (bases 1 to 2432)				
	Simonet,W.S., Lacey,D.L., Dunstan,C.R., Kelley,M., Chang,M.S.,				
	Luthy,R., Nguyen,H.O., Wooden,S., Bennett,L., Boone,T.,				
	Shimamoto,G., DeRose,M., Elliott,R., Colombero,A., Tan,H.L.,				
	Trail,G., Sullivan,J., Davy,E., Bucay,N., Renshaw-Gegg,L.,				
	Hughes,T.M., Hill,D., Pattison,W., Campbell,P., Sander,S., Van,G.,				
	Tarpley,J., Derby,P., Lee,R., Amgen EST Program and Boyle,W.J.				
	Osteoprotegerin: a novel secreted protein involved in the				
	regulation of bone density				
TITLE	Cell 89 (2), 309-319 (1997)				
JOURNAL					
MEDLINE	97262071				


```

SOURCE      Mus musculus (strain:NIH Swiss) cell_line:NIH3T3 CDNA to mRNA.
ORGANISM    Mus musculus
REFERENCE   1 (sites)
AUTHORS     Mizuno,A., Murakami,A., Nakagawa,N., Yasuda,H., Tsuda,E.,
             Morinaga,T. and Higashio,K.
TITLE       Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
             gene and its expression in embryogenesis
JOURNAL     Gene 215 (2), 339-343 (1998)
MEDLINE     98382527
REFERENCE   2 (bases 1 to 2818)
AUTHORS     Mizuno,A. and Yasuda,H.
TITLE       Direct Submission
JOURNAL     Submitted (13-MAY-1998) to the DDBJ/EMBL/GenBank databases. Atsuko
             Mizuno, Snow Brand Milk Products Co., Ltd., Research Institute of
             Life Science; 519 Shimo-Ishibashi, Ishibashi-machi, Tochigi
             329-0512, Japan (E-mail:fvd7042@mb.infoweb.or.jp,
             Tel:0285-52-1331, Fax:0285-53-1314)
COMMENT     On May 20, 1998 this sequence version replaced gi:3135221.
FEATURES    Location/Qualifiers
             1..2818
             /organism="Mus musculus"
             /strain="NIH Swiss"
             /db_xref="taxon:10090"
             /cell_line="NIH3T3"
             238..1443
             /note="mOCIF"
             /codon_start=1
             /product="osteoclastogenesis inhibitory factor"
             /db_xref="PID:d1029209"
             /db_xref="PID:g3135222"
             /translation="MNKMLCCALLVLDIETWTOETLPPKVLHYDPETHGQLLCDKC
             APGYLKQHCRTVRITCLVPCPDHSDYSDSWHTSDCYVCSVPCKELQXVKQCNRTN
             RYCEEGRIEIEFLKRRSGPGGVQAGTPPEPTVCKKCPDGFSGSTSKAPC
             RKHTDCSFGLLLIQGNATHDNCSGNREATQKCGIDVTLCEAAFFFAVPTKTPN
             WLSVLDPGLPTKVNAESVERIKRRHSQEQTFOLLKWKHQRQEMVKKIQDIDL
             CESSVORHGHANLATEOLRALMESLPGKKISPEIERTRKTKCKSEQLLKLLSLWRI
             KNGDDTLKGLMYALKHLKTSHPKTVTHSLRKTFRFLHSFTMYRLYQKLFLEMIGNQ
             VQSVKISCL"
BASE COUNT  844 a 603 c 611 g 760 t
ORIGIN
Query Match      16.1%; Score 74; DB 13; Length 2818;
Best Local Similarity 51.9%; Pred. No. 1.6e-06;
Matches 167; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 16 tactgcaacgtctctgcggggagcggtgagagagagagagagcggtctgccaagccacac 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 481 TATTGACGCCAGTGTGCAAGGAACGTCAGTCGCGTGAAGCAGAGTGCAACCCACCCAC 540
QY 76 aacgtgctgcgcgtgcgcgacacggcttcttcgcgaacgctggttctgcttgagcac 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 541 AACCGAGTGTGTGAGTGTGAGGAGGCGGTACCTGAGATCGAATCTGCTTGAAGCAC 600
QY 136 gcatcgtgccactggtgcgcgcgagtgattgcccgggacccccagccagacagcag 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 601 CGAGAGTGTCCTCCGGGCTCCGGCGGTGCGAAGCTGGAACCCAGACCAACACAGTT 660
QY 196 tgcacgcgtgccccagcactctctcagccagcagcagcagcagcagcagcagtgccag 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 661 TGC AAAAATGTCAGATGGTTCCTCAGGTGAGACITCATCGAAAGCACCCCTGTAGA 720
QY 256 cccacgcgaactgcagcgcgcctggcctggcctcaatgtgccagcttctctccccc 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 721 AAACACACGAGCTGCAGCACATTTGGCTCTCTGCTAATTTCAGAAAGGAATGCAACACAT 780
QY 316 gacacctgtgacacagctgca 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 781 GACACGCTGTGTTCCGGAACA 802

```

```

RESULT      4
LOCUS       AB002146
DEFINITION Homo sapiens mRNA for osteoclastogenesis inhibitory factor (OCIF),
             complete cds.
ACCESSION   AB002146
NID         g3107916
KEYWORDS    Homo sapiens lung fibroblast cell_line:IMR-90 cDNA to mRNA,
             clone_lib:lambda ZAP EXPRESS cDNA library of IMR-90 cells
SOURCE      clone:lambda Olf10.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Yasuda,H., Shima,N., Nakagawa,N., Mochizuki,S., Yano,K., Fujise,N.,
             Sato,Y., Goto,M., Yamaguchi,K., Kuriyama,M., Kanno,T., Murakami,A.,
             Tsuda,E., Morinaga,T. and Higashio,K.
TITLE       Identity of osteoclastogenesis inhibitory factor (OCIF) and
             osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
             osteoclastogenesis in vitro
JOURNAL     Endocrinology 139 (3), 1329-1337 (1998)
MEDLINE     98151033
REFERENCE   2 (bases 1 to 1206)
AUTHORS     Yasuda,H.
TITLE       Direct Submission
JOURNAL     Submitted (25-MAR-1997) to the DDBJ/EMBL/GenBank databases.
             Hisataka Yasuda, Snow Brand Milk Products Co., Ltd., Research
             Institute of Life Science; 519 Shimo-Ishibashi, Ishibashi-machi,
             Tochigi 329-05, Japan (E-mail:fvd7042@mb.infoweb.or.jp,
             Tel:0285-52-1331, Fax:0285-53-1314)
COMMENT     Sequence updated (28-Apr-1998).
FEATURES    Location/Qualifiers
             1..1206
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /cell_line="IMR-90"
             /cell_type="fibroblast"
             /clone_lib="lambda Olf10"
             /clone_lib="lambda ZAP EXPRESS cDNA library of IMR-90
             cells"
             /tissue_type="lung"
sig_peptide 1..63
             /note="signal peptide (aa -21 to -1)"
             /evidence=experimental
CDS         1..1206
             /codon_start=1
             /product="osteoclastogenesis inhibitory factor (OCIF)"
             /db_xref="PID:d1026850"
             /db_xref="PID:g3107917"
             /translation="MNKLLCCALVFLDISIKWTTQETPPPKYLHYDEETSHQLLCDKC
             PPGTYLKQHCRTAKWKTCAPCPDHYTDSWHTSDCYVCSVPCKELQXVKQCNRTN
             RYCECKEGRIEIEFLKRRSGPGGVQAGTPEPTVCKKCPDGFSGSTSKAPC
             RKHTNCSVFGLLITOKGNATHDNCISGNESTQKCGIDVTLCEAAFFFAVPTKTPN
             WLSVLVDNLPGTKVNAESVERIKRHSQEQTFOLLKWKHQRQDQDVLVKKIQDIDL
             CENSQVRIHGHNLTFFEOLRSIMESLPGKKYGAEDIEKTIKACKSDQLLKLLSLWRI
             KNGDDTLKGLMHALKHKSHTHFPAKTVTQSLKKTIRFLHSFTMYRLYQKLFLEMIGNQ
             VQSVKISCL"
             9
variation    9
             /replace="c"
             22
             /replace="t"
             64..1203
             /note="mature OCIF (aa 1 to 380)"
             /evidence=experimental
             /product="osteoclastogenesis inhibitory factor (OCIF)"
             67..189
             /note="cysteine-rich domain 1"
             190..318
             /note="cysteine-rich domain 2"
             319..429
             /note="cysteine-rich domain 3"
             430..558

```


AUTHORS Yasuda,H.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-1997) to the DDBJ/EMBL/GenBank databases.
Hisataka Yasuda, Snow Brand Milk Products Co., Ltd., Research
Institute of Life Science; 519 Shimo-ishibashi, Ishibashi-machi,
Tochigi 329-05, Japan (E-mail:fvb07042@mb.infoweb.or.jp,
Tel:0285-52-1331, Fax:0285-53-1314)
COMMENT Sequence updated (28-Apr-1998).
FEATURES Location/Qualifiers
source
1..1206
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="IMR-90"
/cell_type="fibroblast"
/clone="lambda OIF10"
/clone_lib="lambda ZAP EXPRESS cdna library of IMR-90
cells"
/tissue_type="lung"
sig_peptide
1..63
/note="signal peptide (aa -21 to -1)"
evidence=experimental
CDS
1..1206
/codon_start=1
/product="osteoclastogenesis inhibitory factor (OCIF)"
/db_xref="PID:d1026850"
/db_xref="PID:g3107917"
/translation="MNKLLCCALVFLDISIKWTTQETPPPKYLHYDEETSHQLLCDKC
PPGYLKQHCRAKWKVCAPGDHYDTDSWHSDECLYCSPVKELQIVKQECNRTHN
RVCEKGEYLEIEFCLKHSRCPGFGVVOAGTPERNTVCRCPPDFFSNETSSKAPC
RKHTNCSYFGLLLTKGNATHDNICSGNSESTQKCGIDVTICEAFPAVPTKFTPN
WLSVLNDPGTKNAESVERIKRQHSSEQOTFOLLKWKHONKDDIVKIIQIDIL
CENSVRQIHGANTFEQLRSLMESLPGKVGAEDEIKTIRACKPSDQILKLSLWRI
KNGDDTLKGLMHALKHSKTYHFPKTVTQSLKKTIRFLHSTMYKLYQKLFLEMIGNQ
VQSVKISCL"
9
/replace="c"
variation
22
/replace="t"
mat_peptide
64..1203
/note="mature OCIF (aa 1 to 380)"
evidence=experimental
misc_feature
67..189
/product="osteoclastogenesis inhibitory factor (OCIF)"
/note="cysteine-rich domain 1"
misc_feature
190..318
/note="cysteine-rich domain 2"
misc_feature
319..429
/note="cysteine-rich domain 3"
misc_feature
430..558
/note="cysteine-rich domain 4"
repeat_unit
592..816
/rpt_type=direct
misc_feature
625..858
/note="death domain homologous region 1"
repeat_unit
817..1041
/rpt_type=direct
misc_feature
850..1083
/note="death domain homologous region 2"
BASE COUNT 388 a 283 c 270 g 265 t
ORIGIN
Query Match 14.9%; Score 68.6; DB 41; Length 1206;
Best Local Similarity 50.5%; Pred. No. 2.4e-05;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
Qy 10 tccgcgtactgcaagctctctcgaggagcgtgaggagagcagcggttgcacgcc 69
Db 238 TCTCTATCTAGTCCAGCCCGGTGTCAGGAGCTGCTACGTACGTCAGCAGGAGTGCATCGC 297
Qy 70 acccaaacgctgctgcgcgtgcgcagccggtcttcttcgcagcgtggttctgcttg 129
Db 298 ACCCACACCGCGTGTGCGAATGCAAGGAGGCGCTACTCTGAGTAGAGTTCGCTTG 357

Qy 130 gagcagcatcgtgtccacctggtgcgcgctgattgccccgggcacccccagcagcaac 189
Db 358 AACATAGAGAGTGCCTCTCTGATTTGGATGGTGGCAAGCTGGACCCAGCGGAAT 417
Qy 190 acgcagtgccagcgtgccccccagcaccctcttcagccagcagctccagctcagagcag 249
Db 418 ACAGTTTGCAAAAAGATGTCAGATGGTGTCTTCTCAAAATGAGACGCTCATCTAAAGCACCC 477
Qy 250 tgcagccccaccgcaactgcaagcgccttggccttggccttggccttggccttggccttgc 309
Db 478 TGTAGAAAACACACAAATTCAGATGCTCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 537
Qy 310 tccctgacacccctgtgcaccagctgcactg 340
Db 538 ACACAGCAGACATATGTTCCGAAACAGTG 568
RESULT 7
HSD94332
LOCUS HSD94332 1356 bp mRNA PRI 06-MAY-1997
DEFINITION Human osteoprotegerin (OPG) mRNA, complete cds.
ACCESSION U94332
NID g2072184
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE
1 (bases 1 to 1356)
AUTHORS Simonet,W.S., Lacey,D.L., Dunstan,C.R., Kelley,M., Chang,M.S.,
Luthy,R., Nguyen,H.Q., Wooden,S., Bennett,L., Boone,T.,
Shimamoto,G., DeRose,M., Elliott,R., Colombero,A., Tan,H.I.,
Trail,G., Sullivan,J., Davy,E., Bucay,N., Renshaw-Gegg,L.,
Hughes,T.M., Hill,D., Pattison,W., Campbell,P., Sander,S., Van,G.,
Tarpley,J., Derby,P., Lee,R., Amgen EST Program and Boyle,W.J.
TITLE Osteoprotegerin: a novel secreted protein involved in the
regulation of bone density
JOURNAL Cell 89 (2), 309-319 (1997)
MEDLINE 97262071
REFERENCE 2 (bases 1 to 1356)
AUTHORS Boyle,W.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1997) Department of Cell Biology, Amgen, Inc.,
1840 Dehavilland Drive, Thousand Oaks, CA 91320, USA
FEATURES Location/Qualifiers
source
1..1356
/organism="Homo sapiens"
/db_xref="taxon:9606"
95..1300
/gene="OPG"
95..1300
/gene="OPG"
/codon_start=1
/product="osteoprotegerin"
/db_xref="PID:g2072185"
/translation="MNKLLCCALVFLDISIKWTTQETPPPKYLHYDEETSHQLLCDKC
PPGYLKQHCRAKWKVCAPGDHYDTDSWHSDECLYCSPVKELQIVKQECNRTHN
RVCEKGEYLEIEFCLKHSRCPGFGVVOAGTPERNTVCRCPPDFFSNETSSKAPC
RKHTNCSYFGLLLTKGNATHDNICSGNSESTQKCGIDVTICEAFPAVPTKFTPN
WLSVLNDPGTKNAESVERIKRQHSSEQOTFOLLKWKHONKQADIVKIIQIDIL
CENSVRQIHGANTFEQLRSLMESLPGKVGAEDEIKTIRACKPSDQILKLSLWRI
KNGDDTLKGLMHALKHSKTYHFPKTVTQSLKKTIRFLHSTMYKLYQKLFLEMIGNQ
VQSVKISCL"
BASE COUNT 421 a 326 c 314 g 294 t 1 others
ORIGIN
Query Match 14.9%; Score 68.6; DB 41; Length 1356;
Best Local Similarity 50.5%; Pred. No. 2.4e-05;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```
QY 10 tgcgcactgcaacgtctctcggggagcgtgaggagagcagcagcgttgcacgcc 69
Db 332 TGTCTACTGACCCCGTGTGCAAGAGCTGCAGTACGTACGACGAGGAGTGCATCGC 391
QY 70 acccaacacgctgctcgcctgcgcacccgcttcttcgcacgctgtttctctgtg 129
Db 392 ACCCACAAACCGGTGCGAATGCAAGGAGGGGGCTACCTTGAGATAGAGTCTTGCTG 451
QY 130 gacacgcatgctgcacactgtgcccggcggtgattcccccgggaccccccagcaga 189
Db 452 AAACATAGGAGCTGCCCTCTGATTTGGATTTGGTGAAGCTGGAACCCAGAGCGAAAT 511
QY 190 acgcaagtgcagcgcgtccccccagcgcactctctcagccagcagcctccagtcag 249
Db 512 ACAGTTGCAAAAGATGTCAGATGGGTTCTTCTCAATGAGAGCTCATCTAAGACCC 571
QY 250 tgcagcccccacgcactgcagcgcctggcctgcctcaatgtgcagcgtcttctcc 309
Db 572 TGTAGAAAACACACAATTCAGTGTCTTTGGTCTCTCTGCTAACTCAGAAAGGAATGCA 631
QY 310 tccatgacaccctgtgcacagctgcactg 340
Db 632 ACACAGCAACATATGTTCCGGAACAGTG 662

RESULT 8
MMTNR2A MMTNR2A 1388 bp RNA ROD 17-FEB-1997
LOCUS M.musculus tumor necrosis factor receptor 2 mRNA.
DEFINITION X76401
ACCESSION 9433830
NID 9433830
KEYWORDS tumour necrosis factor receptor.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 1388)
AUTHORS Powell,E.E.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1993) E.E. Powell, c/o John Todd, Level 6
Nuffield Dept of Surgery, The John Radcliffe Hospital, Oxford, UK
REFERENCE 2 (bases 1 to 1388)
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Amino acid variation in the tumor Necrosis factor receptor 2 is
linked to autoimmune diabetes in NOD mice
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1388)
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Allelic variation of the type 2 tumor necrosis factor receptor gene
Mamm. Genome 5 (11), 726-727 (1994)
MEDLINE 95178848
FEATURES
source Location/Qualifiers
1..1388
/organism="Mus musculus"
/strain="NOD"
/db_xref="taxon:10090"
/chromosome="4 (distal region)"
1..1380
/gene="murine tumour necrosis factor receptor 2"
<1..1380
/gene="murine tumour necrosis factor receptor 2"
/codon_start=1
/db_xref="PID:9433831"
/translation="LWATGHTVPAQVLPYKPEGYECOISQEYYDRKAQMCCKACP
PQGYVKNFNTSDTYVCADEASMTQVWQNFRTCLSSGSDTDVETRACTKQNR
VCACEARCYALKTHSGRCQMRKSCGPGFVAGSRAPNGVNLKACAPTFSDTT
STDVCRPHRICISILAPGNASTDVAPESTPLSAIPRTLYVSOPEPTPSQPDQEP
GPSQTSILTSLSGTPILISQSTKGLSLPIGLIVGTSIGLMLGLVNGFILVQRKK
PSCLODKAVPHVPDKSDAVGLEQHLHTAPSSSSSSLESSASAGDRRAPPGGPH
QARVMAEGSSARASISDSHSGHTHVNTCIVNVCSSSDHSSQCSQASATV
GPDFAKPSAPKDEQYFPFSQECPCSPSYETTETTLQSHKEKPLPLGLVPMGMKPSQAGW
FDQIAVKVA"
```

```
variation
260
/gene="murine tumour necrosis factor receptor 2"
/notes="Ser to Thr"
/replace="c"
278
/gene="murine tumour necrosis factor receptor 2"
/notes="Thr to Ile"
/replace="t"
489
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
802
/gene="murine tumour necrosis factor receptor 2"
/notes="Phe to Ile"
/replace="a"
921
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="c"
975
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="c"
1034
/gene="murine tumour necrosis factor receptor 2"
/notes="Ser to Phe"
/replace="t"
1047
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
1143
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
1173
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="a"
1262
/gene="murine tumour necrosis factor receptor 2"
/notes="Thr to Cys"
/replace="g"
1317
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="g"

BASE COUNT 327 a 417 c 371 g 273 t
ORIGIN
```

```
Query Match 11.5%; Score 52.8; DB 13; Length 1388;
Best Local Similarity 52.2%; Pred. No. 0.043;
Matches 117; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 124 tgccttgagcagcatgctgtccacctgtgcggcggtgattgcccccgggacccccagc 183
Db 370 TGCATGAGGCTGAGCAAGTGCGGCCCTGGCTTCGGAGTGGCCAGTTCAGAGGCCCAAT 429
QY 184 cagaacacgcagtcgcagcgcgtgccccccagcgcacttctcagccagcagctccagctca 243
Db 430 GGAATGTGCTATGCAAGGCTGTGCCCCAGGACGTTCTCTGACACCATCATCCACA 489
QY 244 gacagtcgcagcccccacgcacgcacgcgcctggcctggcctccactcaatgtgcaggc 303
Db 490 GATGTGTGACAGGCCACCCGCATCTGTAGCATCTTGGCTATTCCCGGAAATGCAAGCACA 549
QY 304 tctctctcccatgacacctgtgcacagctgcactggtcttccc 347
Db 550 GATGCAGTCTGTGCGCCGAGTCCCAACTCTTAAGTCCCATCCC 593
```

```
RESULT 9
MUSTNFR1 3796 bp mRNA ROD 01-AUG-1991
LOCUS Murine tumor necrosis factor I receptor (TNFR-1) mRNA, complete
DEFINITION cds
ACCESSION M59378
NID g202094
KEYWORDS tumor necrosis factor receptor.
SOURCE Mus musculus lymphoid cDNA to mRNA.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 3796)
AUTHORS Goodwin,R.G., Anderson,D.M., Jerzy,R., Davis,T., Brannan,C.I.,
Copoland,N.G., Jenkins,N.A. and Smith,C.A.
TITLE Molecular cloning and expression of the type 1 and type 2 murine
receptors for tumor necrosis factor
JOURNAL Mol. Cell. Biol. 11, 3020-3026 (1991)
MEDLINE 91246168
FEATURES
source Location/Qualifiers
1..3796
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="7B9"
/cell_type="helper T-cell"
/tissue_type="lymphoid"
/map="Chromosome 4"
43..108
/gene="TNFR-1"
43..1467
/gene="TNFR-1"
/product="tumor necrosis factor receptor 1"
/db_xref="PID:g202095"
/translation="MAPAALWALVLFELQWLWAGHTVPAQVILTPYKPEGYECQISQ
EYDRAQMCACKPPQYVYKFCNKTSDTVACDCEASMTQVWNOFRTCLSGSSSCT
TDQVEIRACTKQNRVACACAGYKALKTSHGSCROCRLSKGPGFGVASSRAPNGN
VLCKACAPGTFSDTTSIDVCRPHRICSLAIPGNASTDAVCAPEPTLSAIPRTLIV
SQPEPTRSQPLDQEPGSPQTSILTSIGSTPIIEQTKGSGISLPIGLIVGTVSLGLM
LGLVNCILVQRKKKPCSLORDAKVPHVPEKSDQAVGLEQQHLLTAPSSSSSLES
SDHSSOSSQASATVGDPAKPSAPKDEQVFPFSECEPQSQCPTTETLQSHKPLP
LGVPDMGMPKPSQAGWFQIAKVA"
43..1467
/gene="TNFR-1"
109..1464
/gene="TNFR-1"
/product="tumor necrosis factor receptor 1"
BASE COUNT 863 a 1052 c 981 g 900 t
ORIGIN Chromosome 4.
Query Match 11.2%; Score 51.2; DB 13; Length 3796;
Best Local Similarity 51.8%; Pred. No. 0.075;
Matches 116; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 124 tgcctggagcagcgcgcgtgtccaccctgtgcccgcgtgattgccccgggcacccccagc 183
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 TGCATGAGGCTGAGCAAGTGGCGCCCTGGCTTCGGAGTGGCCAGTTCAAGAGCCCCCAAT 516
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184 cagacacgcagtcgcagccgtgccccccagcgaccttctcagcagcagctccagctca 243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 GGAATGTGCTATGACAGGCCCTGTGCCCCAGGAGCTTCTGACACCCACCATCATCCACT 576
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244 gagcagtgccagcccccacacgcgcctgcccggcctggccctcaatgtgccaggc 303
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 GATGTGTGAGCGCCCGACCGCATCTGTAGCATCTGGCTATTCGGGAAATGCAAGCACA 636
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 tcttcctcccatgacacctgtgacacagctgcacctgacctgacctcc 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 GATCGAGTCTGTGGCGCGGAGTCCCACTCTAAGTGCCATCCC 680
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
MUSTNFR2 1505 bp mRNA ROD 18-APR-1991
LOCUS Mouse tumor necrosis factor receptor 2 mRNA, complete cds.
DEFINITION
ACCESSION M60469
NID g199827
KEYWORDS transmembrane protein; tumor necrosis factor receptor.
SOURCE Mouse adult macrophage, cDNA to mRNA.
ORGANISM Murinae; Mus.
REFERENCE 1 (bases 1 to 1505)
AUTHORS Lewis,M., Tartaglia,L.A., Lee,A.L., Bennett,G.L., Rice,G.C.,
Wong,G.H.W., Chen,E.Y. and Goeddel,D.V.
TITLE Cloning and expression of cDNAs for two distinct murine necrosis
factor receptors demonstrate one receptor is species specific
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834 (1991)
MEDLINE 91187885
FEATURES
source Location/Qualifiers
1..1505
/organism="Mus musculus"
/db_xref="taxon:10090"
41..106
/sig_peptide
41..1465
CDS
/codon_start=1
/product="murine tumor necrosis factor receptor 2"
/db_xref="PID:g199828"
/translation="MAPAALWALVLFELQWLWAGHTVPAQVILTPYKPEGYECQISQ
EYDRAQMCACKPPQYVYKFCNKTSDTVACDCEASMTQVWNOFRTCLSGSSSCT
TDQVEIRACTKQNRVACACAGYKALKTSHGSCROCRLSKGPGFGVASSRAPNGN
VLCKACAPGTFSDTTSIDVCRPHRICSLAIPGNASTDAVCAPEPTLSAIPRTLIV
SQPEPTRSQPLDQEPGSPQTSILTSIGSTPIIEQTKGSGISLPIGLIVGTVSLGLM
LGLVNCILVQRKKKPCSLORDAKVPHVPEKSDQAVGLEQQHLLTAPSSSSSLES
SDHSSOSSQASATVGDPAKPSAPKDEQVFPFSECEPQSQCPTTETLQSHKPLP
LGVPDMGMPKPSQAGWFQIAKVA"
347 a 459 c 408 g 291 t
ORIGIN
Query Match 11.2%; Score 51.2; DB 13; Length 1505;
Best Local Similarity 51.8%; Pred. No. 0.09;
Matches 116; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 124 tgcctggagcagcgcgcgtgtccaccctgtgcccgcgtgattgccccgggcacccccagc 183
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 TGCATGAGGCTGAGCAAGTGGCGCCCTGGCTTCGGAGTGGCCAGTTCAAGAGCCCCCAAT 514
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184 cagacacgcagtcgcagccgtgccccccagcgaccttctcagcagcagctccagctca 243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 GGAATGTGCTATGCAAGGCCCTGTGCCCCAGGAGCTTCTGACACCCACCATCATCCACT 574
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244 gagcagtgccagcccccacacgcgcctgcccggcctggccctcaatgtgccaggc 303
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 GATGTGTGAGCGCCCGACCGCATCTGTAGCATCTTGGCTATTCGGGAAATGCAAGCACA 634
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 tcttcctcccatgacacctgtgacacagctgcacctgacctgacctcc 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 GATCGAGTCTGTGGCGCGGAGTCCCACTCTAAGTGCCATCCC 678
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AF019046
LOCUS AF019046 2116 bp mRNA ROD 22-NOV-1997
DEFINITION Mus musculus receptor activator of nuclear factor kappa B (RANK)
mRNA, complete cds.
ACCESSION AF019046
NID g2612919
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
```



```

FT protein 22..399 /note="Mature OCIF-CL"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33172.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 62; Page 117-119; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CL in which amino acids 379-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 399 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 399;
Best Local Similarity 41.5%; Pred. No. 8.8e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYNVLCGEREEERACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYSPVKELQYVKQECNRTNRYVCEKGRYLEIEFCLKHSRCPGFGVYQACTPE 137
QY 62 QNTQCPCPPGTFSSASSSEOCOPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNTVCKRCPDGGFSTNETSSKAPCKKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194
QY 122 PGAE--CERAVIDF 134
Db 195 CGIDVILCEEAFFR 209

RESULT 8
R99943
ID R99943 standard; Protein; 351 AA.
AC R99943;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CC.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT protein 22..351
FT /note= "Mature OCIF-CC"

PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33173.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 65; Page 119-121; 183pp; Japanese.

```

```

CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CC in which amino acids 331-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 351 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 351;
Best Local Similarity 41.5%; Pred. No. 7.6e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYNVLCGEREEERACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYSPVKELQYVKQECNRTNRYVCEKGRYLEIEFCLKHSRCPGFGVYQACTPE 137
QY 62 QNTQCPCPPGTFSSASSSEOCOPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNTVCKRCPDGGFSTNETSSKAPCKKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194
QY 122 PGAE--CERAVIDF 134
Db 195 CGIDVILCEEAFFR 209

RESULT 9
R99944
ID R99944 standard; Protein; 272 AA.
AC R99944;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CDD2.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT protein 22..272
FT /note= "Mature OCIF-CDD2"

PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33174.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 68; Page 121-122; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CDD2 in which amino acids 252-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 272 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 272;

```


Best Local Similarity 41.5%; Pred. No. 5.8e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVVKQECNRTHNRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 10
R99948
ID R99948 standard; Protein; 393 AA.
AC R99948;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CBst
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Synthetic.
FH Key
FT peptide 1. .21 Location/Qualifiers
FT /note= "Signal peptide"
FT protein 22. .393
FT /note= "Mature OCIF-CBst"
FT misc_difference 392
FT /label= Gln371Leu
FT W09626217-A1.
PN 29-AUG-1996.
PD 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33178.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 80; Page 126-128; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.
CC These changes are caused by the introduction of a restriction site in
CC the DNA encoding this protein. The OCIF of the invention has a
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 393 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 393;
Best Local Similarity 41.5%; Pred. No. 8.7e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVVKQECNRTHNRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 10
R99948
ID R99948 standard; Protein; 393 AA.
AC R99948;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CBst
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Synthetic.
FH Key
FT peptide 1. .21 Location/Qualifiers
FT /note= "Signal peptide"
FT protein 22. .393
FT /note= "Mature OCIF-CBst"
FT misc_difference 392
FT /label= Gln371Leu
FT W09626217-A1.
PN 29-AUG-1996.
PD 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33178.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 80; Page 126-128; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.
CC These changes are caused by the introduction of a restriction site in
CC the DNA encoding this protein. The OCIF of the invention has a
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 393 AA;

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 11
R99933
ID R99933 standard; Protein; 401 AA.
AC R99933;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C21S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Synthetic.
FH Key
FT peptide 1. .21 Location/Qualifiers
FT /note= "Signal peptide"
FT protein 22. .401
FT /note= "Mature OCIF-C21S"
FT misc_difference 277
FT /label= C21S
FT W09626217-A1.
PN 29-AUG-1996.
PD 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33163.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 35; Page 98-100; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C21S in which the 21st Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVVKQECNRTHNRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 12
R99934
ID R99934 standard; Protein; 401 AA.
AC R99934;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C22S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.

```

OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..401
FT misc_difference 277
FT /note= "Mature OCIF-C22S"
FT /label= C22S
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DELCYSPVKELQYVKQECNRTNRNVCCKEGRYLEIEFCLKHRSCTPPGFGVQAGTPE 137
OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db 138 RNTVCKRCPDGFFSNSTSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194
OY 122 PGAE--CERAVIDF 134
Db 195 CGIDVTLCCEAERFR 209

RESULT 13
R99949
ID R99949 standard; Protein; 321 AA.
AC R99949; 1997 (first entry)
DE Mutated OCIF, OCIF-CSph.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..321
FT /note= "Mature OCIF-CSph"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DELCYSPVKELQYVKQECNRTNRNVCCKEGRYLEIEFCLKHRSCTPPGFGVQAGTPE 137
OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db 138 RNTVCKRCPDGFFSNSTSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194
OY 122 PGAE--CERAVIDF 134
Db 195 CGIDVTLCCEAERFR 209

RESULT 13
R99949
ID R99949 standard; Protein; 321 AA.
AC R99949; 1997 (first entry)
DE Mutated OCIF, OCIF-CSph.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..321
FT /note= "Mature OCIF-CSph"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

```

```

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T33179.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 83; Page 128-129; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CSph in which amino acids 298-380 of the mature
CC OCIF protein are replaced by Ser-leu-Asp. These changes are caused by
CC the introduction of a restriction site in the DNA encoding this protein.
CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD
CC under reducing conditions and 120 kD under non-reducing conditions. The
CC protein is adsorbed onto cation-exchangers or heparin and its activity is
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost
CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone
CC resorption and therefore in the treatment and prevention of disorders
CC of bone resorption, e.g. osteoporosis.
SQ Sequence 321 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 321;
Best Local Similarity 41.5%; Pred. No. 6.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DELCYSPVKELQYVKQECNRTNRNVCCKEGRYLEIEFCLKHRSCTPPGFGVQAGTPE 137
OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db 138 RNTVCKRCPDGFFSNSTSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 194
OY 122 PGAE--CERAVIDF 134
Db 195 CGIDVTLCCEAERFR 209

RESULT 14
R99936
ID R99936 standard; Protein; 360 AA.
AC R99936; 1997 (first entry)
DE Mutated OCIF, OCIF-DCRI.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..360
FT /note= "Mature OCIF-DCRI"
FT misc_difference 22..23
FT /note= "Position of deletion, delta 2-42"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T33166.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 44; Page 105-107; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-DCRI in which amino acids 2-42 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

```

CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 360 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 360;
Best Local Similarity 41.5%; Pred. No. 7.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

Qy 2 ERCRYCNVLCGEREEARACHATHNRACRTGTFFAHAGFCLHASCPPGAGVIAPGTPS 61
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 37 DECLYCSPVKRELQYVKQECNRTHNRVCEKRGYLEIEFCLKHKRSCPPGFGVQAGTPE 96
:
Qy 62 QNTQCPCPPGTFSSASSSQCPHRCNTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 97 RNTVCKRCPDGFFSNETSAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 153
:
Qy 122 PGAE--CERAVIDF 134
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 154 CGIDVTLCEEAFFRF 168

RESULT 15
R99924
ID R99924 standard; Protein; 380 AA.
AC R99924;
DT 22-APR-1997 (first entry)
DE Mature osteoclastogenesis inhibitory factor.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS osteoporosis.
QS Homo sapiens.
PN WO9628217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI; 96-402320/40.
DR N-PSDB; T36685.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 6; Page 62-64; 183pp; Japanese.
CC This sequence represents the mature osteoclastogenesis inhibitory
CC factor (OCIF) of the invention. The OCIF has a molecular weight by
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
CC reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
CC useful in the control of bone resorption and therefore in the
CC treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 380 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 380;
Best Local Similarity 41.5%; Pred. No. 8.3e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

Qy 2 ERCRYCNVLCGEREEARACHATHNRACRTGTFFAHAGFCLHASCPPGAGVIAPGTPS 61
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 57 DECLYCSPVKRELQYVKQECNRTHNRVCEKRGYLEIEFCLKHKRSCPPGFGVQAGTPE 116
:
Qy 62 QNTQCPCPPGTFSSASSSQCPHRCNTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 117 RNTVCKRCPDGFFSNETSAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSESTQK 173
:
Qy 122 PGAE--CERAVIDF 134
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | |

Db 174 CGIDVTLCEEAFFRF 188

Search completed: May 13, 1999, 17:21:17
Job time: 20716 sec

This Page Blank (uspto)

Result No.	Score	Query %			DB	ID	Description
		Match	Length	Length			
1	841	100.0	300	1	W63622	Human tumour necro	
2	841	100.0	300	1	W68102	Amino acid sequenc	
3	337	40.1	170	1	W63623	Human tumour necro	
4	324.5	38.6	326	1	R99940	Mutated OC1F, OC1F	
5	311.5	37.0	401	1	R99931	Mutated OC1F, OC1F	
6	309.5	36.8	327	1	R99941	Mutated OC1F, OC1F	
7	309.5	36.8	399	1	R99942	Mutated OC1F, OC1F	
8	309.5	36.8	351	1	R99943	Mutated OC1F, OC1F	
9	309.5	36.8	351	1	R99943	Mutated OC1F, OC1F	
10	309.5	36.8	272	1	R99944	Mutated OC1F, OC1F	
11	309.5	36.8	393	1	R99948	Mutated OC1F, OC1F	
12	309.5	36.8	401	1	R99933	Mutated OC1F, OC1F	
13	309.5	36.8	401	1	R99934	Mutated OC1F, OC1F	
14	309.5	36.8	321	1	R99949	Mutated OC1F, OC1F	
15	309.5	36.8	360	1	R99936	Mutated OC1F, OC1F	
16	309.5	36.8	380	1	R99924	Mature osteoclasto	
17	309.5	36.8	401	1	R99925	Full length osteoc	
18	309.5	36.8	390	1	R99357	Human tumour necro	
19	309.5	36.8	401	1	W83845	Human osteoprotege	
20	309.5	36.8	391	1	W53238	Human OC1F genome	
21	309.5	36.8	401	1	W53239	Human OC1F genome	
22	306.5	36.4	401	1	W57636	Modified TR1 recep	
23	305.5	36.3	401	1	W57635	Rat osteoprotegeri	
24	303.5	36.1	401	1	W83844	TR1 receptor prote	
25	302.5	36.0	401	1	R99932	Mouse osteoprotege	
26	302	35.9	197	1	R99945	Mutated OC1F, OC1F	
27	302	35.9	187	1	R99950	Mutated OC1F, OC1F	
28	300.5	35.7	401	1	R99935	Mutated OC1F, OC1F	
29	243	28.9	461	1	R11001	40kD TNF inhibitor	
30	243	28.9	392	1	R11605	Human 75kD TNF-bi	
31	243	28.9	235	1	W52270	Tumour necrosis fa	
32	243	28.9	235	1	W59665	Human soluble tumo	
33	241	28.7	461	1	R11141	Human TNF-R deduce	
34	241	28.7	485	1	R24016	Fusion protein TNF	
35	241	28.7	461	1	R24058	Fibroblast derived	
36	241	28.7	461	1	R72504	p75 Tumour Necrosi	
37	241	28.7	359	1	R99937	Mutated OC1F, OC1F	
38	237	28.2	518	1	R51003	Sequence of a reco	
39	232	27.6	183	1	R77421	BamTP delc53 nerv	
40	218.5	26.0	474	1	R11422	TNF-R deduced from	
41	213.5	25.6	461	1	R51002	Sequence of human	
42	205.5	24.4	359	1	R99939	Mutated OC1F, OC1F	
43	193	22.9	143	1	R99946	Mutated OC1F, OC1F	

RESULT	2
W66102	
ID	W66102 standard; Protein; 300 AA.
AC	W66102;
DT	02-DEC-1998 (first entry)

DE Amino acid sequence of tumour necrosis related receptor (TR4).
 KW Human: tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition: chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 OS Homo sapiens.
 PN EP-861850-A1.
 PD 02-SEP-1998.
 PR 20-JAN-1998; 300382.
 PF 04-FEB-1997; US-794796.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Emery J, Tan KB, Truneh A, Young PR;
 DR WPI: 98-508248/44.
 DR N-PSDB: V07654.
 PT New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 PS Claim 1; Fig 1; 21pp; English.
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 SQ Sequence 300 AA;

Query Match 100.0%; Score 841; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 3.4e-66;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LERCRCYVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTP 60
 |||||
 DB 85 LERCRCYVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTP 144
 |||||

QY 61 SONTQOCPCPGTFSASSSSQCOQHNRCTALGLNVPSSSHDTLCTSCGTGFFPLSTR 120
 |||||
 DB 145 SONTQOCPCPGTFSASSSSQCOQHNRCTALGLNVPSSSHDTLCTSCGTGFFPLSTR 204
 |||||

QY 121 VPGAECERAVIDFVAFQDISIRKRLQLOALE 153
 |||||
 DB 205 VPGAECERAVIDFVAFQDISIRKRLQLOALE 237
 |||||

RESULT 3
 W63623
 ID W63623 standard; Protein; 170 AA.
 AC W63623;
 DT 26-OCT-1998 (first entry)
 DE Human tumour necrosis factor receptor-6 beta protein.
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..170
 FT note="TNFR-6 beta"
 FT Region 31..166
 FT /note="Soluble extracellular domain"
 PN W09830694-A2.
 PD 16-JUL-1998.
 PE 13-JAN-1998; U00153.
 PR 14-JAN-1997; US-035496.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 DR WPI: 98-399142/34.
 DR N-PSDB: V39086.
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)

PS Claim 20; Fig 2A-2B; 91pp; English.
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 SQ Sequence 170 AA;

Query Match 40.1%; Score 337; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1.4e-22;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LERCRCYVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPG 58
 |||||
 DB 85 LERCRCYVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPG 142
 |||||

RESULT 4
 R99940
 ID R99940 standard; Protein; 326 AA.
 AC R99940;
 DT 23-APR-1997 (first entry)
 DE Mutated OCIF, OCIF-DDD1.
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW osteoporosis.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT protein 22..326
 FT /note="Signal peptide"
 FT /note="Mature OCIF-DDD1"
 FT misc_difference 198..199
 FT /note="Position of deletion, delta 178-252"
 PN W09626217-A1.
 PD 29-AUG-1996.
 PF 20-FEB-1996; J00374.
 PR 20-FEB-1995; JP-054977.
 PR 21-JUL-1995; JP-207508.
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 DR WPI: 96-402320/40.
 DR N-PSDB: T33170.
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 PS Claim 56; Page 113-114; 183pp; Japanese.
 CC This sequence represents a mutated version of the full length
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
 CC sequence represents OCIF-DDD1 in which amino acids 178-252 of the
 CC mature OCIF protein are deleted. The OCIF of the invention
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
 CC cation-exchangers or heparin and its activity is lowered after 10 mins
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
 CC deg.C. OCIF is useful in the control of bone resorption and therefore
 CC in the treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.
 SQ Sequence 326 AA;

Query Match 38.6%; Score 324.5; DB 1; Length 326;
 Best Local Similarity 37.9%; Pred. No. 3.5e-21;
 Matches 58; Conservative 33; Mismatches 57; Indels 5; Gaps 2;

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:03 ; Search time 33.52 seconds
(without alignments)
170.985 Million cell updates/sec

Title: US-09-212-270-2
Perfect score: 841
Sequence: 1 LERCRCYNLCGEREEARA.....FVAFQDISIKRLQLLALE 153

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR_58:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	241	28.7	461	2 A35356	tumor necrosis fac
2	221.5	26.3	459	2 I48854	gene murine tumour
3	218.5	26.0	474	2 B38634	tumor necrosis fac
4	176.5	21.0	349	2 D36858	gene G4R protein -
5	176	20.9	435	2 I54182	tumor necrosis fac
6	157.5	18.7	325	2 B43592	T2 protein - rabbl
7	154.5	18.4	277	2 A60771	B-cell activation
8	149.5	17.8	277	2 I37552	Ox40 homologue - h
9	148	17.6	289	2 A46515	B cell-associated
10	148	17.6	305	2 A6476	CD40 - mouse
11	145.5	17.3	271	2 S12783	OX40 antigen precu
12	140.5	16.7	416	2 JN0006	nerve growth facto
13	139.5	16.6	595	2 A42086	CD30 antigen precu
14	137.5	16.3	272	2 I48700	gene ox40 protein
15	131	15.6	425	2 A26431	nerve growth facto
16	123.5	14.7	326	1 GOV2ML	T2 protein - myxom
17	123.5	14.7	427	1 GOUJUN	nerve growth facto
18	123	14.6	256	2 B32393	T-cell antigen 4-1
19	117	13.9	255	2 I38426	4-1BB - human
20	116	13.8	255	2 JT0752	lymphocyte activat
21	106.5	12.7	461	1 GQRTT1	tumor necrosis fac
22	105.5	12.5	327	2 A46484	apoptosis-mediati
23	102.5	12.2	454	1 GQKST1	tumor necrosis fac
24	102.5	12.2	454	2 I57826	tumor necrosis fac
25	102	12.1	780	2 A34102	von Willebrand fac
26	102	12.1	335	2 A40036	apoptosis-mediati
27	101.5	12.1	899	2 G02428	prohormone convert
28	101.5	12.1	915	2 JC6148	subtilisin-like pr
29	101.5	12.1	2321	2 S78549	Notch3 protein - h
30	101	12.0	969	2 A39490	serine proteinase
31	101	12.0	975	2 JC5570	subtilisin-like pr
32	101	12.0	962	2 JCS571	subtilisin-like pr
33	100	11.9	642	2 S52111	uromodulin precurs
34	99.5	11.8	1607	1 MMWSB2	laminin gamma-1 ch
35	98.5	11.7	328	2 S00530	surface antigen CR
36	98	11.7	1680	2 A34334	furin (EC 3.4.21.7
37	97.5	11.6	1609	1 MMHUB2	laminin gamma-1 ch
38	97	11.5	314	2 I37383	FAS soluble protei
39	96.5	11.5	915	2 B48225	probable proteol

40	96	11.4	2482	2	A47648	perlecan homolog u
41	96	11.4	2277	2	B47648	perlecan homolog u
42	96	11.4	1160	2	C47648	perlecan homolog u
43	96	11.4	1713	2	A55347	adhesive ligand ep
44	95.5	11.4	915	2	A48225	subtilisin-like pr
45	95.5	11.4	1548	2	S34583	serine proteinase

ALIGNMENTS

RESULT 1

A35356
tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C>Date: 14-Sep-1990 #sequence_revision 14-Sep-1990 #text_change 20-Mar-1998
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:Cross-references: GB:M32315; NID:gl89185; PID:gl89186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195, 'R', 197-461 <KOH>
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PID:g339758
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular
A:Reference number: A48416; MUID:91370690
A:Accession: A48416
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-461 <DEM>
A:Cross-references: GB:S63368; NID:g235648; PID:g235649
A:Note: Sequence extracted from NCBI backbone (NCBIN:63368, NCBIP:63371)
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons
A:Reference number: A36007; MUID:90349572
A:Accession: A36007
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>
A:Cross-references: GB:M35857; NID:g339751; PID:g339752
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, J. Biol. Chem. 265, 20131-20138, 1990
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor
A:Reference number: A23666; MUID:91056048
A:Accession: A23666
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-40; 65-69; 136-141; 300-306 <LOE>
R:Engelmann, H.; Novick, D.; Wallach, D. J. Biol. Chem. 265, 1531-1536, 1990
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Eviden
A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D. Gene 150, 381-386, 1994

```

A:Reference number: I38094; MUID:95121934
A:Accession: I38094
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: EMBL:X80021; NID:g666044; PID:g825701
C:Genetics:
A:Gene: GDB:TNFR2
A:Cross-references: GDB:125914; OMIM:191191
A:Map position: ip36.2-1p36.2
A:Introns: 26/3
A:Note: the list of introns is incomplete
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>
F:40-76/Domain: NGF receptor repeat homology <NG1>
F:78-119/Domain: NGF receptor repeat homology <NG2>
F:120-162/Domain: NGF receptor repeat homology <NG3>
F:164-201/Domain: NGF receptor repeat homology <NG4>
F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>
F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      28.7%; Score 241; DB 2; Length 461;
Best Local Similarity 36.2%; Pred. No. 1.5e-13;
Matches 46; Conservative 20; Mismatches 49; Indels 12; Gaps 3;

QY      4  CRYCNVLGGEREEEARACHATHNACRCRTGFFAHAG-----FCLEHACPPGAGVIAP 57
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      93  CLCSGSRCSQDVETQACTREQNRICTCPGWYKALSKQEGCRICAPLKRCPGFGVARP 152
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      58  GTPSQNTQOCPPCPGTFSSASSSEQOCPHNCRTALGALNVPGSSSHDTLCTCTGFPPL 117
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      153  GTFESDVVCRPCAGCTGTFSTSSIDICRPHQICNVVA----IPGNASMDAVCTSTST--PT 206
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      118  STRVPGA 124
      : | | |
Db      207  RSMAPGA 213

RESULT      2
I48854
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Feb-1997
C:Accession: I48854
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A:Reference number: I48854; MUID:95178848
A:Accession: I48854
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X76401; NID:g433830; PID:g433831
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

Query Match      26.3%; Score 221.5; DB 2; Length 459;
Best Local Similarity 37.2%; Pred. No. 6.9e-12;
Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

QY      4  CRYCNVLGGEREEEARACHATHNACRCRTGFF----AHAGF---CLEHACPPGAGVIA 56
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      79  CLCSGSSCSTDVETRACTKQNRVCAACEAGRYCALKTHSGSCRCQRLSKCGFGFGVAS 138
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      57  PGTSPNTQOCPPCPGTFSSASSSEQOCPHNCRTALGALNVPGSSSHDPLC 109
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      139  SRAPNGNVLKACAGPFTSDTSTDVCRPHRICSIILA----IPGNASTDAVC 187

```


A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H P
 A:Reference number: S4688
 A:Accession: S4688
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <KOL>
 A:Cross-references: EMBL:X67117; NID:g516428; PID:g516449
 A:Experimental source: strain India-1967, isolate Ind3
 R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
 FEBS Lett. 319, 80-83, 1993
 A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protective
 A:Reference number: S32385
 A:Accession: S32385
 A:Molecule type: DNA
 A:Residues: 31-168 <SHC>
 A:Cross-references: EMBL:X69198
 A:Experimental source: strain India-1967, ssp. major
 C:Genetics:
 A:Gene: G4R
 C:Superfamily: NGF receptor repeat homology
 F:68-109/Domain: NGF receptor repeat homology <NG2>
 F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match 21.0%; Score 176.5; DB 2; Length 349;
 Best Local Similarity 31.3%; Pred. No. 3.5e-08;
 Matches 41; Conservative 21; Mismatches 62; Indels 7; Gaps 2;
 Qy 1 LERCYCNVLCGEREEARACHATHNRACRTGTGFF-----AHAGFCLHASCPPGAGV 54
 Db 80 LPACLSGRCGRNSQVETRSCNTTNIRICECPGYCLLKSGSGKACVSTQKCGIGYV 139
 Qy 55 IAPGPPSNTQOCPPPTGFSASSSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTG 114
 Db 140 -SGHTSVGDVCSGCFGTYSHTVSADKCEVPNNTFNIDVETLIPVNDTSCRTTT 198
 Qy 115 PPLSTRVPGAE 125
 Db 199 TGLSESLTSE 209

RESULT 5
 I54182
 tumor necrosis factor receptor 2-related protein - human
 C:Species: Homo sapiens (man)
 C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 29-Aug-1997
 C:Accession: I54182
 R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
 Genomics 16, 214-218, 1993
 A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed sequen
 A:Reference number: I54182; MUID:93252381
 A:Accession: I54182
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-435 <RES>
 A:Cross-references: GB:L04270; NID:g339761; PID:g339762
 C:Genetics:
 A:Gene: GDB:LTBR
 A:Cross-references: GDB:1230195; OMIM:600979
 A:Map position: 12p13.3-12p13.1

Query Match 20.9%; Score 176; DB 2; Length 435;
 Best Local Similarity 34.9%; Pred. No. 4.7e-08;
 Matches 44; Conservative 12; Mismatches 56; Indels 14; Gaps 6;

Qy 4 CRYCNVLCGEREEARACHATHNRACRTGFFAHAGFCLC--H---ASCPPGA-GVIA 56
 Db 101 CRCPDPVNG--LEETAPCTSKRKTQRCQPGNFC-AAWALETHCELSDCPGTEAEIK 157
 Qy 57 PGTPSNTQOCPPPTGFSASSSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTGFP 116
 Db 158 DEVGKGNHCVPKAGHFQNTSSPARGCPHRCENQGLVEAAPGTAQSDTTCKN---P 213

Qy 117 LSTRVP 122
 Db 214 LEPLPP 219
 RESULT 6
 B43692
 T2 protein - rabbit fibroma virus
 C:Species: rabbit fibroma virus, Shope fibroma virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-May-1997
 C:Accession: B43692
 R:Upton, C.; DeLange, A.M.; McFadden, G.
 Virology 160, 20-30, 1987
 A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomer
 A:Reference number: A43692
 A:Accession: B43692
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <UPT>
 A:Cross-references: GB:M17433
 C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
 F:64-105/Domain: NGF receptor repeat homology <NG2>
 F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 18.7%; Score 157.5; DB 2; Length 325;
 Best Local Similarity 31.6%; Pred. No. 1.3e-06;
 Matches 37; Conservative 16; Mismatches 53; Indels 11; Gaps 3;
 Qy 4 CRYCNVLCGEREEARACHATHNRACRTGTGFFA-----HAGFCLHASCPPGAGVIAP 57
 Db 79 CVSCRGCTGHLSESQPCDRTDRVCNCTGNYCLLKQNGCRICAPQTKCPAGYGV-SG 137
 Qy 58 GTPSNTQOCPPPTGFSASSSSSQCPHNRCTALGLALNVPGSSSHDTLCTCTG 114
 Db 138 HTRAGDTLCEKCPHTYSDLSLSPTRCSTFNYSVGFNL----YPVNETSCTTTAG 190

RESULT 7
 A60771
 B-cell activation protein CD40 precursor - human
 N:Alternate names: B-cell surface antigen Bp50
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 10-Sep-1997
 C:Accession: S04460; A60771
 R:Stamenkovic, I.; Clark, E.A.; Seed, B.
 EMBO J. 8, 1403-1410, 1989
 A:Title: A B-lymphocyte activation molecule related to the nerve growth factor recept
 A:Reference number: S04460; MUID:89356608
 A:Accession: S04460
 A:Molecule type: mRNA
 A:Residues: 1-277 <STA>
 A:Cross-references: EMBL:X60592; NID:g29850; PID:g29851
 R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.
 J. Immunol. 142, 562-567, 1989
 A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-
 A:Reference number: A60771
 A:Accession: A60771
 A:Molecule type: protein
 A:Residues: 21-50 <BRA>
 A:Experimental source: Burkitt lymphoma cell line Raj1
 C:Genetics:
 A:Gene: GDB:CD40
 A:Cross-references: GDB:215268; OMIM:109535
 A:Map position: 20q12-20q13.2
 C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-277/Product: B-cell activation protein CD40 #status experimental <NAT>
 F:21-193/Domain: extracellular #status predicted <EXT>
 F:194-215/Domain: transmembrane #status predicted <TMM>
 F:216-277/Domain: intracellular #status predicted <CYT>
 F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Query Match
Best Local Similarity 18.4%; Score 154.5; DB 2; Length 277;
Matches 34; Conservative 17; Mismatches 52; Indels 5; Gaps 2;

QY 5 RYCNVLCGEREEARACHATHNRACRCRTGFFAHA---GFCLEHASCPPGAGVIAPGTPS 61
DB 81 KYCDPNLGLRVQOKGTSER--DTICICEGWHCTSAECSCVLHRSCTSPGFGVKQIATGV 138

QY 62 QNTQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 139 SDTICEPCPVGFPSNVSSAFKCHPTWCETKDLVVOQAGTKNTDVC 186

RESULT 8
OX40 homologue - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I37552
R:Latza, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fonat
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homologue: cDNA structure, expression and chromosomal assignment o
A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:g472957; PID:g472958

Query Match
Best Local Similarity 17.8%; Score 149.5; DB 2; Length 277;
Matches 35; Conservative 9; Mismatches 39; Indels 23; Gaps 4;

QY 4 RYCNVLCGEREEARACHATHNRACRCRTGFFAHAAGFLEHASCPPGAGVIAPGTPSQN 63
DB 84 CWCNLRSG--SERKOLCFATQTVCRKAG-----TQPLDSYKPG-----123

QY 64 TQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 123 VDCAPCPGPHF--SPGDNQACKPWTNCTLAGKHTLQPA NSSDAIC 166

RESULT 9
A46515
B cell-associated surface molecule CD40 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne,
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; MUID:93094586
A:Accession: A46515
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Experimental source: BALB/c, liver
A>Note: sequence extracted from NCBI backbone (NCBIP:120357)

Query Match
Best Local Similarity 17.6%; Score 148; DB 2; Length 289;
Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

QY 3 RC---RYCNVLCGEREEARACHATHNRACRCRTGFFA---HAGFLEHASCPPGAGVIA 56
DB 76 RCHQHRHCPEHQGLRVKKEGT--AESDTVCTCKEGQHCSTKDCEACAQHTPCIPGFGYME 133

QY 57 PGTFSQNTQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
A>Note: sequence extracted from NCBI backbone (NCBIP:75206, NCBIP:75207)
C:Keywords: transmembrane protein

Query Match
Best Local Similarity 17.6%; Score 148; DB 2; Length 305;
Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

QY 3 RC---RYCNVLCGEREEARACHATHNRACRCRTGFFA---HAGFLEHASCPPGAGVIA 56
DB 76 RCHQHRHCPEHQGLRVKKEGT--AESDTVCTCKEGQHCSTKDCEACAQHTPCIPGFGYME 133

QY 57 PGTFSQNTQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 134 MATETDTVCHPCVPVGFPSNVSSAFKCHPTWCETKDLVVOQAGTKNTDVC 186

RESULT 11
SI2783
OX40 antigen precursor - rat
N:Alternate names: nerve growth factor receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
C:Accession: SI2783; S08036
R:Mallett, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphoc
A:Reference number: SI2783; MUID:90214614
A:Accession: SI2783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:g57830; PID:g57831
C:Keywords: growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-271/Product: OX40 antigen #status predicted <MAT>
F:211-235/Domain: transmembrane #status predicted <TMM>

Query Match
Best Local Similarity 17.3%; Score 145.5; DB 2; Length 271;
Matches 36; Conservative 13; Mismatches 31; Indels 35; Gaps 5;

QY 2 ERCRYCNVLCGEREEARACHATHNRACRCRTGFFAAGFLEHASCPPGAGVIAPGT-P 60
DB 77 KOCTQCNHRSG--SELKONCTPTEDTVCCOR-----PGTQP 110

QY 61 SONT-----QCQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 111 RQDSSHKLGVDCVPCPPGHF--SPGSNQACKPWTNCTLSGKQIRHPASNSLDTVC_163

RESULT 12
JN0006
nerve growth factor receptor, low affinity - chicken
N:Alternate names: NGF receptor
C:Species: Gallus gallus (Chicken)
```

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 01-May-1998
C:Accession: JN0006; A60504
R:Large, T.H.; Weskamp, G.; Helder, J.C.; Radeke, M.J.; Misko, T.P.; Shooter, E.M.; Reid
Neuron 2, 1123-1134, 1989
A>Title: Structure and developmental expression of the nerve growth factor receptor in t
A:Reference number: JN0006; MUID:90166579
A:Accession: JN0006
A:Molecule type: mRNA
A:Residues: 1-416 <LAR>
A:Experimental source: embryonic chick brain
R:Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.
Dev. Biol. 137, 287-304, 1990
A>Title: Structure and developmental expression of the chicken NGF receptor.
A:Reference number: A60504
A:Accession: A60504
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 21-35, Y', 37-172, 'K', 174-275, 'S', 277-395, 'R', 397-416 <HEU>
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
C:Comment: This protein is thought to form a high-affinity receptor when it associates
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; t
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-416/Product: nerve growth factor receptor #status predicted <MAT>
F:21-239/Domain: extracellular #status predicted <EXT>
F:24-57/Domain: NGF receptor repeat homology <NG1>
F:59-100/Domain: NGF receptor repeat homology <NG2>
F:101-139/Domain: NGF receptor repeat homology <NG3>
F:141-181/Domain: NGF receptor repeat homology <NG4>
F:189-237/Region: serine/threonine-rich
F:240-261/Domain: transmembrane #status predicted <MEM>
F:262-416/Domain: intracellular #status predicted <INT>
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.7%; Score 140.5; DB 2; Length 416;
Best Local Similarity 29.8%; Pred. No. 4.6e-05;
Matches 34; Conservative 16; Mismatches 49; Indels 15; Gaps 4;

Qy 2 ECRYCNVLCGEREERACHATHNRACRRTGTFPAH--AGFCLHASCPPGAGVIAPGT 59
Db 73 EPCKPCTQCVGLHMSA-PCVESDDAVRCAYGFQDELSGSKCSICEVGFGLMFPGR 131

Qy 60 PSQNTQCPGPTGTSASSSSSECCQPHRNCNTALGLALNVPGSSSHDTLCSTCT 113
Db 132 DSQDVTCECPGTSDANFVDCPLP---CTI-----CEENEVNVKECT 173

RESULT 13
A42086
CD30 antigen precursor - human
N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Sep-1998
C:Accession: A42086
R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
Cell 68, 421-427, 1992
A>Title: Molecular cloning and expression of a new member of the nerve growth factor rec
A:Reference number: A42086; MUID:92154659
A:Accession: A42086
A:Molecule type: mRNA
A:Residues: 1-595 <DUR>
A:Cross-references: GB:M83554; NID:g180095; PID:g180096
A:Experimental source: HUT-102 cell line
A>Note: sequence extracted from NCBI backbone (NCBI:82088, NCBI:82090)
C:Genetics:
A:Gene: GDB:CD30; D1S166E
A:Cross-references: GDB:131547; OMIM:153243
A:Map position: 1p36-1p36
C:Superfamily: NGF receptor repeat homology
C:Keywords: glycoprotein; growth factor receptor; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-383/Domain: extracellular #status predicted <EXT>
F:384-407/Domain: transmembrane #status predicted <TM>
F:408-595/Domain: intracellular #status predicted <CYT>
F:101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.6%; Score 139.5; DB 2; Length 595;
Best Local Similarity 31.4%; Pred. No. 7.7e-05;
Matches 38; Conservative 12; Mismatches 50; Indels 21; Gaps 6;

Qy 2 ECRYCNVLCGERE-EEARACHATHNRACRRTGTFPAHGF-----CLEHASCPPGAGV 54
Db 79 DRCTAC-VTCRDDLVEKTPCAWNSRVCEGRMFCSTSAVNSCARCFHSHVCPAGMIV 137
Qy 55 IAPGTPSONTOCQCPPTGTSASSSSSECCQ-----PHRNCNTALGLALNVPGSSSHDT 107
Db 138 KFGPTAQNTVCEPASPGRV-SPACASPCNKPEPSGTIPQAKPTPVS-----PATSSAST 191

Qy 108 L 108
Db 192 M 192

RESULT 14
I48700
gene ox40 protein - mouse
N:Alternate names: OX40 antigen
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996
C:Accession: I48700; I48334; S34377
R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J.
J. Immunol. 151, 5261-5271, 1993
A>Title: Cloning of mouse Ox40: a T cell activation marker that may mediate T-B cell
A:Reference number: I48700; MUID:94044750
A:Accession: I48700
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-272 <RES>
A:Cross-references: EMBL:Z21674; NID:g312827; PID:g312828
R:Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
Eur. J. Immunol. 25, 926-930, 1995
A>Title: Gene structure and chromosomal localization of the mouse homologue of rat OX
A:Reference number: I48334; MUID:95255413
A:Accession: I48334
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-14, 'G', 16-272 <RE2>
A:Cross-references: EMBL:X85214; NID:g732818; PID:g732819
C:Genetics:
A:Gene: ox40
A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

Query Match 16.3%; Score 137.5; DB 2; Length 272;
Best Local Similarity 30.4%; Pred. No. 5.7e-05;
Matches 35; Conservative 11; Mismatches 34; Indels 35; Gaps 5;

Qy 2 ECRYCNVLCGEREERACHATHNRACRRTGTFPAHGFCLHASCPPGAGVIAPGT-P 60
Db 78 KOCTOCNHRSG--SELKQNCPTQDTCRCR-----PGTOP 111

Qy 61 SQNT-----CQCPGPTGTSASSSSSECCQPHRNCNTALGLALNVPGSSSHDTLC 109
Db 112 RDSGYKLGVDVCPGPGHF--SPGNQACKPWTNCTLSGRQTRHPASDLSDAVC 164

RESULT 15
A26431
nerve growth factor receptor precursor, low affinity - rat
N:Alternate names: NGF receptor
C:Species: Rattus norvegicus (Norway rat)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 01-May-1998
C:Accession: A26431; PH1229

R:Radeke, M.J.; Misko, T.P.; Hsu, C.; Herzenberg, L.A.; Shooter, E.M.
 Nature 325, 593-597, 1987
 A:Title: Gene transfer and molecular cloning of the rat nerve growth factor receptor.
 A:Reference number: A26431; MUID:87115859
 A:Accession: A26431
 A:Molecule type: mRNA
 A:Residues: 1-425 <RAD>
 A:Cross-references: GB:X05137; NID:g56755; PID:g56756
 R:Metz, M.; Timmusk, T.; Allikmets, R.; Saarma, M.; Persson, H.
 Gene 121, 247-254, 1992
 A:Title: Regulatory elements and transcriptional regulation by testosterone and retinoic acid.
 A:Reference number: PH1229
 A:Accession: PH1229
 A:Molecule type: DNA
 A:Residues: 1-20 <MET>
 A:Cross-references: GB:X61269
 C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma cells, and on the cysteine-rich region of the extracellular domain may form part or all of a disulfide bridge.
 C:Comment: This protein is thought to form a high-affinity receptor when it associates with a specific ligand.
 C:Genetics:
 A:Introns: 20/3
 C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; tyrosine kinase
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-425/Product: nerve growth factor receptor #status predicted <MAT>
 F:30-251/Domain: extracellular #status predicted <EXT>
 F:33-66/Domain: NGF receptor repeat homology <NG1>
 F:68-109/Domain: NGF receptor repeat homology <NG2>
 F:110-148/Domain: NGF receptor repeat homology <NG3>
 F:150-190/Domain: NGF receptor repeat homology <NG4>
 F:198-249/Region: serine/threonine-rich
 F:252-273/Domain: transmembrane #status predicted <MEM>
 F:274-425/Domain: intracellular #status predicted <INT>
 F:61/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.6%; Score 131; DB 2; Length 425;
 Best Local Similarity 27.3%; Pred. No. 0.0003;
 Matches 36; Conservative 15; Mismatches 53; Indels 28; Gaps 5;
 QY 2 ECRYCNVLCGEREEPARACHATHNRACRTGFF--AHAGFCLHSCPPGAGVIAPGT 59
 Db 82 EPKPCTECLGLQMSA-PCVEADDAVCRCAVGYODEETHGHCACSVCEVSGGLVFSQ 140
 QY 60 PSQNTQCQPCPTFSASSSSSQCPHNCALGLALNVPSSSHDYLCTSC--TGFP 117
 Db 141 DRQNTVCECPGTYSDANHVDPCLP-----CTVCEPTEROL 178
 QY 118 STRVPGAE-ECE 128
 Db 179 RECTPWDAECE 190

Search completed: May 13, 1999, 20:59:04
 Job time: 2318 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 1999, 14:43:36 ; Search time 25.8 seconds
(without alignments)
159.169 Million cell updates/sec

Title: US-09-212-270-2
Perfect score: 841
Sequence: 1 LERCRCYNLVCGEREEARA.....FVAFQDISIKRLQLLQALE 153

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	28.9	461	1	TNR2_HUMAN
2	218.5	26.0	474	1	TNR2_MOUSE
3	176.5	21.0	349	1	VC22_VARV
4	176	20.9	435	1	TNRC_HUMAN
5	165	19.6	415	1	TNRC_MOUSE
6	157.5	18.7	325	1	VT2_SEVKA
7	154.5	18.4	277	1	CD40_HUMAN
8	149.5	17.8	277	1	OX40_HUMAN
9	148	17.6	289	1	CD40_MOUSE
10	145.5	17.3	271	1	OX40_RAT
11	140.5	16.7	416	1	NGFR_CHICK
12	139.5	16.6	595	1	CD30_HUMAN
13	137.5	16.3	269	1	CD40_BOVIN
14	137.5	16.3	272	1	OX40_MOUSE
15	131	15.6	425	1	NGFR_RAT
16	123.5	14.7	427	1	NGFR_HUMAN
17	123.5	14.7	326	1	VT2_MXAVL
18	123	14.6	256	1	41BB_MOUSE
19	117	13.9	255	1	41BB_HUMAN
20	107	12.7	323	1	FASA_BOVIN
21	106.5	12.7	461	1	TNR1_RAT
22	105.5	12.5	327	1	FASA_MOUSE
23	105.5	12.5	417	1	WSL1_MOUSE
24	102.5	12.2	454	1	TNR1_HUMAN
25	102	12.1	335	1	FASA_HUMAN
26	101	12.0	969	1	PAC4_HUMAN
27	99.5	11.8	1607	1	LMG1_MOUSE
28	98.5	11.7	328	1	C170_GIALA
29	98	11.7	1680	1	FUR2_DROME
30	97.5	11.6	1609	1	LMG1_HUMAN
31	97	11.5	1799	1	LMR2_MOUSE
32	96.5	11.5	915	1	PAC6_RAT
33	96	11.4	1713	1	LMR3_HUMAN
34	96	11.4	2481	1	UNS2_CAEEL
35	96	11.4	2813	1	VWF_CANFA
36	95.5	11.4	915	1	PAC6_MOUSE
37	95	11.3	2569	1	LMR3_MOUSE
38	93.5	11.1	816	1	NEL_HUMAN
39	92.5	11.0	3110	1	LMR2_HUMAN
40	92.5	11.0	3707	1	PGBW_MOUSE
41	92.5	11.0	461	1	TNR1_PIG
42	92.5	11.0	640	1	UROM_HUMAN
43	92	10.9	998	1	EPB3_HUMAN

44 91.5 10.9 3075 1 LMA1_HUMAN P25391 homo sapien
45 91 10.8 3635 1 LMA5_MOUSE Q61001 mus musculu

ALIGNMENTS

RESULT 1	
TNR2_HUMAN	
ID TNR2_HUMAN STANDARD; PRT; 461 AA.	
AC P20333;	
DT 01-FEB-1991 (REL. 17, CREATED)	
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)	
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)	
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR	
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B).	
GN TNR2 OR TNFR.	
OS HOMO SAPIENS (HUMAN).	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC EUTHERIA; PRIMATES.	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE; 90260639.	
RA SMITH C.A., DAVIS T., ANDERSON D., SOLAM L., BECKMANN M.P., JERZY R.,	
RA DOWER S.K., COSMAN D., GOODWIN R.G.;	
RL SCIENCE 248:1019-1023(1990).	
RN [2]	
RP SEQUENCE FROM N.A.	
RX MEDLINE; 91045991.	
RA KOHNO T., BREWER M.T., BAKER S.L., SCHWARTZ P.E., KING M.W.,	
RL HALE K.K., SOIRES C.H., THOMPSON R.C., VANNICE J.L.;	
RN [3]	
RP PROC. NATL. ACAD. SCI. U.S.A. 87:8331-8335(1990).	
RN [3]	
RP SEQUENCE FROM N.A.	
RX MEDLINE; 96299745.	
RA BELTINGER C.P., WHITE P.S., MARIS J.M., SULMAN E.P., JENSEN S.J.,	
RA LEPASLIER D., STALLARD B.J., GOEDDEL D.V., DESAUVAGE F.J.,	
RL BRODEUR G.M.;	
RN [4]	
RP GENOMICS 35:94-100(1996).	
RN [4]	
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.	
RX MEDLINE; 90349572.	
RA HELLER R.A., SONG K., ONASCH M.A., FISCHER W.H., CHANG D.,	
RA RINGOLD G.M.;	
RN [5]	
RP PROC. NATL. ACAD. SCI. U.S.A. 87:6151-6155(1990).	
RN [5]	
RP SEQUENCE OF 27-31.	
RX MEDLINE; 90110215.	
RA ENGELMANN H., NOVICK D., WALLACH D.;	
RL J. BIOL. CHEM. 265:1531-1536(1990).	
RN [6]	
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.	
RX MEDLINE; 91056048.	
RA LOETSCHER H., SCHLAEGER E.J., LAHM H.-W., PAN Y.-C.E., LESSLAUER W.,	
RA BROCKHAUS M.;	
RN [7]	
RL J. BIOL. CHEM. 265:20131-20138(1990).	
RN [7]	
RP CHARACTERIZATION.	
RX MEDLINE; 93016040.	
RA PENNICA D., LAM V.T., MIZE N.K., WEBER R.F., LEWIS M., FENDLY B.M.,	
RA LIPARI M.T., GOEDDEL D.V.;	
RN [8]	
RL J. BIOL. CHEM. 267:21172-21178(1992).	
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND	
CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.	
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.	
CC -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW	
CC LEVEL ON THREONINE RESIDUES.	
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.	
DR EMBL; M2315; G189186; -	
DR EMBL; M35857; G339752; -	
DR EMBL; U52165; G1469541; -	
DR EMBL; U52156; G1469541; JOINED.	
DR EMBL; U52157; G1469541; JOINED.	

Db 154 SRAPNGVLKACAPGTFSDTTSTSDVCRPHRICSYLA-----IPGNASTDAVC 202

```
RESULT 3
VC22_VARY
ID VC22_VARY STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS VARIOLA VIRUS.
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE: 93202281.
RA SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;
RL FEBS LETT. 319:80-83(1993).
CC -1- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
DR EMBL: X69198; G457087; -.
DR EMBL: X67117; G516449; -.
DR PIR: D36858; D36858.
DR PIR: S35987; S35987.
DR PIR: S46888; S46888.
DR HSP: P19438; 1TNFR.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW REPEAT.
FT DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MW; 50D0B435 CRC32;
```

Query Match 21.0%; Score 176.5; DB 1; Length 349;
Best Local Similarity 31.3%; Pred. No. 3.1e-08;
Matches 41; Conservative 21; Mismatches 62; Indels 7; Gaps 2;

```
QY 1 LERCRCVNLGGEREEERACHATHNRACRGTGFF-----AHAGFCLEHASCPPGAGV 54
DB 80 LPALCSNGRCNSQVETRSCNTHNRICECSGYLLKGGSGKACVQTKCGIGYGV 139
QY 55 IAPGTSQNTQCPGPGTFSASSSSSEOCQPHRNCNTALGLALNVGSSSHDTLCTSCGT 114
DB 140 -SGHTSVGDVICSFGGTYSHTVSSADKCEPVNNTFNYIDVEITLYPVNDTSCRTTTT 198
QY 115 FPLSTRVPGAE 125
DB 199 TGLSESILTSE 209
```

```
RESULT 4
TNRC_HUMAN
ID TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTBR OR TNFR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 93252381.
RA BAENS M., CHAFFANET M., CASSIMAN J.J., DEN BERGHE H., MARYNEN P.;
RL GENOMICS 16:214-218(1993).
```

```
RN FUNCTION.
RP MEDLINE: 94225209.
RA CROWE P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.,
RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
RL SCIENCE 264:707-710(1994).
CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: L04270; G339762; -.
DR HSP: P19999; 1CLG.
DR MIM: 600979; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 3.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
FT SIGNAL 1 30
FT CHAIN 31 435
FT DOMAIN 31 227 LYMPHOTOXIN-BETA RECEPTOR.
FT TRANSMEM 228 248 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 249 435
FT DOMAIN 42 211 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 81 4 X TNFR-CYS.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 168 TNFR-CYS 3.
FT REPEAT 169 211 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 POTENTIAL.
FT CARBOHYD 177 177 POTENTIAL.
SQ SEQUENCE 435 AA; 46709 MW; 203B82DD CRC32;
```

Query Match 20.9%; Score 176; DB 1; Length 435;
Best Local Similarity 34.9%; Pred. No. 4.2e-08;
Matches 44; Conservative 12; Mismatches 56; Indels 14; Gaps 6;

```
QY 4 CRYCNVLCGEREEERACHATHNRACRGTGFFAHAGFCLE--H----ASCPPGA-GVIA 56
DB 101 CRPCDPVMG--LEEIAPCTSKRKTQRCQPGMFC-AAWALECTHCELLSDCPPGTEALK 157
QY 57 PGTPSQNTQCPGPGTFSASSSSSEOCQPHRNCNTALGLALNVGSSSHDTLCTSCGTGP 116
DB 158 DEVGKGNHCVFKAGHFQNTSSPASCQPHTRCENQGLVEAAPGTAQSDTTCKN----P 213
QY 117 LSTRVP 122
DB 214 LLEPLPP 219
```

```
RESULT 5
TNRC_MOUSE
ID TNRC_MOUSE STANDARD; PRT; 415 AA.
AC P50284;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN LTBR OR TNFR.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CVB; TISSUE-LUNG;
```



```
RA ZHENG Z., NALSMITH J.H., THOMAS D.;
RL PROTEIN SCI. 7:1124-1135(1998).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X60592; E1175755; -.
DR PIR: S04460; S04460.
DR PDB: 1CDF; 01-APR-97.
DR MM: 109535; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS50050; TNFR_NGFR_2; 4.
KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL;
KW 3D-STRUCTURE.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 277 CD40L RECEPTOR.
FT DOMAIN 20 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT DISULFID 26 37
FT DISULFID 38 51
FT DISULFID 41 59
FT DISULFID 62 77
FT DISULFID 83 103
FT DISULFID 105 119
FT DISULFID 111 116
FT DISULFID 125 143
FT CARBOHYD 153 153 POTENTIAL.
FT CARBOHYD 180 180 POTENTIAL.
SQ SEQUENCE 277 AA; 30619 MM; 38284411 CRC32;

Query Match 18.4%; Score 154.5; DB 1; Length 277;
Best Local Similarity 31.5%; Pred. No. 1.7e-06;
Matches 34; Conservative 17; Mismatches 52; Indels 5; Gaps 2;

OY 5 RYCNVLCGEREERERACHATHNRACRCRTGFFAHA---GFCLEHASCPPGAGVIAPGTPS 61
DB 81 KYCDPNLRLRVQOKTSET--DTICTCEGWHCTSEACECVLHRSCTSPGFGVKQIATGV 138
OY 62 QNTCCPCPPGPTFSASSSSSECCOPHRNCTALGLALNVPSSSHDTLC 109
DB 139 SDTICEPCPGTFSNVSSAFECHECHWTSCETKDYVQAGTKTIVVC 186

RESULT 8
OX40_HUMAN STANDARD; PRT: 277 AA.
ID OX40_HUMAN
AC P43489; 013663;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-FRANSCRIPTIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TXGP1L.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94170844.
RA LAZZA U., DUEKROD H., SCHNITZER S., RINGELING J., EITELBACH F.,
RA HUMMEL M., FONATSCH C., STEIN H.;
RL EUR. J. IMMUNOL. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95219871.
RA BAUM P.R., GAILE R.B. III, RAMSDELL F., SRINIVASAN S., SORESENSEN R.A.,
```

```
RA WATSON M.L., SELDIN M.F., CLIFFORD K.N., GRABSTEIN K., ALDERSON M.R.;
RL CIRC. SHOCK 44:30-34(1994).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X75962; G472958; -.
DR EMBL: S76792; G913406; ALT_INIT.
DR MM: 600315; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 277 OX40L RECEPTOR.
FT DOMAIN 29 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 235 POTENTIAL.
FT DOMAIN 236 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 169 4 X TNFR-CYS.
FT REPEAT 30 65 TNFR-CYS 1.
FT REPEAT 66 107 TNFR-CYS 2.
FT REPEAT 108 126 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 127 167 TNFR-CYS 4.
FT CARBOHYD 146 146 POTENTIAL.
FT CARBOHYD 160 160 POTENTIAL.
SQ SEQUENCE 277 AA; 29340 MM; E0EB6765 CRC32;

Query Match 17.8%; Score 149.5; DB 1; Length 277;
Best Local Similarity 33.0%; Pred. No. 4.6e-06;
Matches 35; Conservative 9; Mismatches 39; Indels 23; Gaps 4;

OY 4 CRYCNVLCGEREERERACHATHNRACRCRTGFFAHAAGFCLEHASCPPGAGVIAPGTPSON 63
DB 84 CTWCMILRSG--SERQKLTATQDTVCRCRAG-----TQPLDSYKPG----- 123
OY 64 TCCPCPPGPTFSASSSSSECCOPHRNCTALGLALNVPSSSHDTLC 109
DB 123 VDCACCPGPHF--SPGDNQACKPWTNCTLAGKHTLQPLANSSSDAIC 166

RESULT 9
CD40_MOUSE STANDARD; PRT: 289 AA.
ID CD40_MOUSE
AC P27512;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDM40).
GN CD40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92105763.
RA TORRES R.M., CLARK E.A.;
RL J. IMMUNOL. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C;
RA TORRES R.M.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-LIVER;
RX MEDLINE: 93094586.
RA GRIMALDI J.C., TORRES R., KOZAK C.A., CHANG R., CLARK E.A.,
RA HOWARD M., COCKAYNE D.A.;
RL J. IMMUNOL. 149:3921-3926(1992).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M83312; G1553059; -.

```

DR EMBL: M94126; G192526; JOINED.
 DR EMBL: M94129; G192526; JOINED.
 DR EMBL: M94128; G192526; JOINED.
 DR EMBL: M94127; G192526; JOINED.
 DR PIR: A46476; A46476.
 DR HSP: P19438; 1TNR.
 DR MGJ: M8336; CD40.
 DR PROSITE: PS00652; TNR_NGFR_1; 1.
 DR PROSITE: PS50050; TNR_NGFR_2; 4.
 DR RECEPTOR: B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 289
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 289
 FT DOMAIN 25 187
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT CARBOHYD 153
 FT SEQUENCE 289 AA; 32111 MW; D8D70A2C CRC32;

Query Match 17.6%; Score 148; DB 1; Length 289;
 Best Local Similarity 31.9%; Pred. No. 6.4e-06;
 Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

OY 3 RC---RYCNVLGCEEREERACHATNHRACRCRTGFFA---HAGFCLFNHASCPPGAGVIA 56
 DB 76 RCHQHHCCEPNOGLRVKKEGT--AESDVTCTCKEGHCHTSKDEACAOHPCCPGFVME 133
 OY 57 PGTSPONTCCPCPGTPTFSASSSSSECCOPHRNCTALGLALNVPGSSSHDTLC 109
 DB 134 MATETDTYCHPCVPVGFSSQSLFEKCYPTWCSEDKNLEVLKGTSGTNYVIC 186

RESULT 10
 OX40_RAT STANDARD; PRI; 271 AA.
 ID OX40_RAT
 AC P15725;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
 GN TXGP1L OR OX40.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-T-CELL.
 RX MEDLINE: 90214614.
 RA MALETTE S., FOSSUM S., BARCLAY A.N.;
 RL EMOB J. 9:1063-1068(1990).
 CC -1- FUNCTION: RECEPTOR FOR THE OX40L/SP34 CYTOKINE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: X17037; G57831; .
 DR PIR: S08036; S08036.
 DR PROSITE: PS12783; S12783.
 DR PROSITE: PS00652; TNR_NGFR_1; 3.
 DR PROSITE: PS50050; TNR_NGFR_2; 2.
 DR RECEPTOR: T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
 KW SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 271
 FT DOMAIN 20 210
 FT TRANSMEM 216 235
 FT DOMAIN 236 271
 FT DOMAIN 25 164
 FT REPEAT 60
 FT TNR-CYS 1.
 FT TNR-CYS 1.

FT REPEAT 61 102
 FT REPEAT 103 123
 FT REPEAT 124 164
 FT CARBOHYD 143 143
 FT SEQUENCE 271 AA; 29895 MW; 65460ED2 CRC32;

Query Match 17.3%; Score 145.5; DB 1; Length 271;
 Best Local Similarity 31.3%; Pred. No. 9.7e-06;
 Matches 36; Conservative 13; Mismatches 31; Indels 35; Gaps 5;

OY 2 ERCRYCNVLGCEEREERACHATNHRACRCRTGFFAAGFCLFNHASCPPGAGVIA 60
 DB 77 KQCTQCNHSG--SELKQNTPTEDTVCQCR-----FGTGP 110
 OY 61 SONT-----QCPCPPGTPTFSASSSSSECCOPHRNCTALGLALNVPGSSSHDTLC 109
 DB 111 RQDSHRLGVCVCPGPHF--SPGSNQACKPWTNCTLSGKQIRHPSNSLDTVC 163

RESULT 11
 NGFR_CHICK STANDARD; PRI; 416 AA.
 ID NGFR_CHICK
 AC P18519;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
 DE (GP80-LNGFR) (P75 ICD).
 GN NGFR.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 90166579.
 RA LARGE T.H., WESKAMP G., HEIDER J.C., RADEKE M.J., MISKO T.P.,
 RA SHOOTER E.M., REICHARDT L.F.;
 RL NEURON 2:1123-1134(1989).
 RN [2]
 RP SEQUENCE OF 21-416 FROM N.A.
 RX MEDLINE: 90152140.
 RA HEIDER J.G., FATEME-NAINE S., WHEELER E.F., BOTHWELL M.;
 RL DEV. BIOL. 137:287-304(1990).
 CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
 CC NT-3, AND NT-4.
 CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
 CC BOND FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNR-TYPE CYSTEINE-RICH REGION.
 DR PIR: J00006; J00006.
 DR PIR: A60504; A60504.
 DR PROSITE: PS00652; TNR_NGFR_1; 3.
 DR PROSITE: PS50050; TNR_NGFR_2; 3.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 DR RECEPTOR: NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
 KW PHOSPHORYLATION; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 416
 FT DOMAIN 29 239
 FT TRANSMEM 240 261
 FT DOMAIN 262 416
 FT DOMAIN 23 181
 FT DOMAIN 23 57
 FT REPEAT 58 100
 FT REPEAT 101 140
 FT REPEAT 141 181
 FT REPEAT 188 236
 FT DOMAIN 333 410
 FT DISULFID 24 35
 FT DISULFID 36 49
 FT BY SIMILARITY.

```
FT DISULFID 39 56 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 78 91 BY SIMILARITY.
FT DISULFID 81 99 BY SIMILARITY.
FT DISULFID 101 114 BY SIMILARITY.
FT DISULFID 117 130 BY SIMILARITY.
FT DISULFID 120 138 BY SIMILARITY.
FT DISULFID 141 156 BY SIMILARITY.
FT DISULFID 159 172 BY SIMILARITY.
FT DISULFID 162 180 BY SIMILARITY.
FT CARBOHYD 52 52 POTENTIAL.
FT CONFLICT 36 36 C -> Y (IN REF. 2).
FT CONFLICT 173 173 T -> K (IN REF. 2).
FT CONFLICT 276 276 N -> S (IN REF. 2).
FT CONFLICT 396 396 K -> R (IN REF. 2).
SQ SEQUENCE 416 AA: 44654 MW: 4D3F086A CRC32;

Query Match 16.7%; Score 140.5; DB 1; Length 416;
Best Local Similarity 29.8%; Pred. No. 3.8e-05;
Matches 34; Conservative 16; Mismatches 49; Indels 15; Gaps 4;

QY 2 ERRCYCNVLGGEREERARACHATHNRACRCRTGFFAH--AGFCLERHASCPPGAGVIAPGT 59
Db 73 EPCPKPTQCVGLHSMRA-PCVESDDAVRCACAYCFODELSGCKECSICGVGFIAPPCR 131
QY 60 PSONTCQCPQCPPTGFSASSSSSEOCOPHRNCTALGLALNVPGSSSHDTICTSCT 113
Db 132 DSQDTVCECEPCEGTFSEAFNFDPLP---CTI-----CEENFVMVKECT 173

RESULT 12
CD30_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CD30L RECEPTOR PRECURSOR (LYMPHOCYTE ACTIVATION ANTIGEN CD30) (KI-1 ANTIGEN).
GN CD30.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOID;
RX MEDLINE: 92154659.
RA DUEKOP H., LAITZA U., HUMMEL M., EITELBACH F., SEED B., STEIN H.;
RL CELL 68:421-427(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95089787.
RA JUNG W., KRUEGER S., RENNER C., GAUSE A., SAHIN U., TRUMPER L.,
RA PEREUNSCHEU M.;
RL MOL. IMMUNOL. 31:1329-1334(1994).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD30L. MAY PLAY
CC A ROLE IN THE REGULATION OF CELLULAR GROWTH AND TRANSFORMATION OF
CC ACTIVATED LYMPHOBLASTS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DISASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD30 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/ProW/cd/cd30.htm".
DR EMBL: M83554; G180096; -.
DR EMBL: S75768; E184151; -.
DR PIR: A42086; A42086.
DR MIM: 153243; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS00500; TNFR_NGFR_2; 2.
KW T-CELL; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL; RECEPTOR; REPEAT.
FT SIGNAL 1 18
FT CHAIN 19 595 CD30L RECEPTOR.
```

```
FT DOMAIN 19 379 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 380 407 POTENTIAL.
FT DOMAIN 408 595 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 325 6 X TNFR-CYS.
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT REPEAT 205 241 TNFR-CYS 4.
FT REPEAT 243 281 TNFR-CYS 5.
FT REPEAT 282 325 TNFR-CYS 6.
FT DOMAIN 347 377 PRO/SER/THR-RICH.
FT CARBOHYD 101 101 POTENTIAL.
FT CARBOHYD 276 276 POTENTIAL.
SQ SEQUENCE 595 AA: 63747 MW: A46D6C0B CRC32;

Query Match 16.6%; Score 139.5; DB 1; Length 595;
Best Local Similarity 31.4%; Pred. No. 6.4e-05;
Matches 38; Conservative 12; Mismatches 50; Indels 21; Gaps 6;

QY 2 ERRCYCNVLGGERE-DEARACHATHNRACRCRTGFFAHAGF-----CLEHASCPPGAGV 54
Db 79 DRCTAC-VTCSRDLYEKTPCANSSRYVCECRGMCSISAVNSCARCFHSVCPAGMIV 137
QY 55 IAPGTSQNTQCPQCPPTGFSASSSSSEOC-----PHRNCTALGLALNVPGSSSHDT 107
Db 138 KFPGTAKNTKVCBPASPGV-SPACASPEKCEKPSGTTIPQAKPTPV-----PATSSAST 191
QY 108 L 108
Db 192 M 192

RESULT 13
CD40_BOVIN STANDARD; PRT; 269 AA.
ID CD40_BOVIN
AC Q28203;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (FRAGMENT).
GN CD40.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA ESTES D.M., HIRANO A.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: U57745; G1480643; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS00500; TNFR_NGFR_2; 1.
KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 269
FT DOMAIN 20 193 CD40L RECEPTOR.
FT TRANSMEM 194 215 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 216 269 POTENTIAL.
FT DOMAIN 25 187 CYTOPLASMIC (POTENTIAL).
FT REPEAT 25 60 4 X TNFR-CYS.
FT REPEAT 61 103 TNFR-CYS 1.
FT REPEAT 104 144 TNFR-CYS 2.
FT REPEAT 145 187 TNFR-CYS 3.
FT CARBOHYD 153 153 TNFR-CYS 4.
FT CARBOHYD 180 180 POTENTIAL.
FT NON_TER 269
SQ SEQUENCE 269 AA: 29983 MW: 46C3AB4D CRC32;

Query Match 16.3%; Score 137.5; DB 1; Length 269;
```

Best Local Similarity 31.5%; Pred. No. 4.5e-05;
Matches 35; Conservative 11; Mismatches 54; Indels 11; Gaps 2;

OY 5 RYCVVIGEREEREA-----RACHATHNRACRCRTGFAHAGFCLEHASCPRGAGVIAPG 58
DB 81 RYCNPNJGLRLOSGLTNTDTICVCEGQHCSTHT-----CECTPHSLCLPGGVKQIA 135
OY 59 TSPONTOCPCPCPTGTFSSSSSEOCOPHRNCITAGLALNPVSSSHDTLC 109
DB 136 TGLDTCVCEPCLGFFSNVSSAFECHEKHWTCERKGLVEQHVGNKTIYVC 166

RESULT 14
OX40_MOUSE STANDARD; PRT: 272 AA.
ID OX40_MOUSE
AC PA7741;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TXGP1 OR OX40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE: 94044750.
RA CALDERHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J.,
RA CLAASSEN E., NOELLE R.J., FELL H.;
RL J. IMMUNOL. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95255413.
RA BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA BARCLAY A.N.;
RL EUR. J. IMMUNOL. 25:926-930(1995).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: Z21674; G312828; -;
DR EMBL: X85214; G732819; -;
DR MGD: MGI:104512; TXGP1.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 272
FT DOMAIN 20 211
FT TRANSMEM 212 236
FT DOMAIN 237 272
FT DOMAIN 26 165
FT REPEAT 26 61
FT REPEAT 62 103
FT REPEAT 104 124
FT REPEAT 125 165
FT CARBOHYD 144 144
FT CONFLICT 15 15
SQ SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

Query Match 16.3%; Score 137.5; DB 1; Length 272;
Best Local Similarity 30.4%; Pred. No. 4.6e-05;
Matches 35; Conservative 11; Mismatches 34; Indels 35; Gaps 5;

OY 2 ECRATCNVLGEREEERACHATHNRACRCRTGFAHAGFCLEHASCPRGAGVIAPG-T-P 60
DB 78 KQCTQCNRHSG--SELKNCNCTPTDTCRCR-----PGTGP 111
OY 61 SONT-----OCQPCPCPTGTFSSSSSEOCOPHRNCITAGLALNPVSSSHDTLC 109
DB 112 RODSGYKLGVDVCPGPHF--SPGNNOACKPMTNCTLSGQTRHPADSIDAVC 164

RESULT 15

NGFR_RAT STANDARD; PRT: 425 AA.
ID NGFR_RAT
AC P07174;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE (GP80-LINGFR) (P75 ICD).
NGFR.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87115859.
RA RADEKE M.J., MISKO T.P., HSU C., HERZENBERG L.A., SHOOTER E.M.;
RL NATURE 325:593-597(1987).
RN [2]
RP SEQUENCE OF 1-22 FROM N.A.
RC TISSUE-LIVER.
RX MEDLINE: 93077038.
RA METSIS M., TIMMUSK T., ALLIKMETS R., SARMA M., PERSSON H.;
RL GENE 121:247-254(1992).
RN [3]
RP STRUCTURE BY NMR OF 334-418.
RX MEDLINE: 97449145.
RA LIEPINSH E., ILAG L.L., OTTING G., IBANEZ C.F.;
RL EMBO J. 16:4999-5005(1997).
CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
NT-3, AND NT-4.
CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
BOND FORMATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X05137; G65756; -; NOT_ANNOTATED_CDS.
DR PIR: A26431; A26431.
DR PDB: INGR; 29-JUL-97.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 4.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
KW PHOSPHORYLATION; SIGNAL; 3D-STRUCTURE.
FT SIGNAL 1 29
FT CHAIN 30 425
FT DOMAIN 30 251
FT TRANSMEM 252 273
FT DOMAIN 274 425
FT DOMAIN 32 190
FT REPEAT 32 66
FT REPEAT 67 108
FT REPEAT 109 148
FT REPEAT 149 190
FT DOMAIN 198 249
FT DOMAIN 354 419
FT DISULFID 33 44
FT DISULFID 45 58
FT DISULFID 48 65
FT DISULFID 68 84
FT DISULFID 87 100
FT DISULFID 110 123
FT DISULFID 126 139
FT DISULFID 129 147
FT DISULFID 150 165
FT DISULFID 168 181
FT DISULFID 171 189
FT CARBOHYD 71 71
FT POTENTIAL.

This Page Blank (uspto)


```

ID 070202 PRELIMINARY: PRT: 401 AA.
AC 070202:
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE OSTEOCLASTOGENESIS INHIBITORY FACTOR.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIH SWISS;
RA MIJUNO A., MURAKAMI A., NAKAGAWA N., YASUDA H., TSUDA E., MORINAGA T.,
RA HIGASHIO K.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129/OLA;
RX MEDLINE: 98382527.
RA MIJUNO A., MURAKAMI A., NAKAGAWA N., YASUDA H., TSUDA E., MORINAGA T.,
RA HIGASHIO K.;
RL HIGASHIO K.;
RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
gene and its expression in embryogenesis.";
RL GENE 215:339-343(1998).
DR EMBL: AB013898; D1029209; -.
DR EMBL: AB013903; D1034354; -.
DR EMBL: AB013899; D1034354; JOINED.
DR EMBL: AB013900; D1034354; JOINED.
DR EMBL: AB013901; D1034354; JOINED.
DR EMBL: AB013902; D1034354; JOINED.
SQ SEQUENCE 401 AA; 45935 MW; 944D4023 CRC32;

```

```

Query Match 35.7%; Score 300.5; DB 10; Length 401;
Best Local Similarity 37.9%; Pred. No. 1.9e-15;
Matches 58; Conservative 25; Mismatches 65; Indels 5; Gaps 2;

```

```

QY 2 ECRRCNVLCGEREEARACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIA 61
DB DECVYCSFVCKELQSVKCECNTHNRVCECEGRLEIEFCIKHNSCPGSGVQAGTPE 137
QY 62 ONTOCOPCPGPTGTFSSASSSSSEOCOPHRNCTALGIALNVPSSSHDTLCSTGFP 121
DB 138 PNTVCKKCPDGFSESTSKAPCRKHTDCSTFGLLIOKGNATHNV---CGSNEATOK 194
QY 122 PGAEE--CERAVIDEFAFODISIKRLQLLQAL 152
DB 195 CGIDVTLCEAEFFRAVPTKILIPNMLSVLYDSL 227

```

```

RESULT 6
Q16042 PRELIMINARY: PRT: 425 AA.
AC 016042:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMIINAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 91370690.
RA DEMBIC Z., LOETSCHER H., GUBLER U., PAN Y.C., LAHM H.W., GENTZ R.,
RA BROCKHAUS M., LESSLAUER W.;
RT "Two human TNF receptors have similar extracellular, but distinct
intracellular, domain sequences.";
RL CYTOKINE 2:231-237(1990).
DR EMBL: S63368; E91326; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PFAM: PF00020; TNFR_c6; 4.
SQ SEQUENCE 425 AA; 44608 MW; A6B87BF7 CRC32;

```

```

Query Match 28.7%; Score 241; DB 2; Length 425;
Best Local Similarity 36.2%; Pred. No. 5.1e-11;
Matches 46; Conservative 20; Mismatches 49; Indels 12; Gaps 3;
QY 4 CRYCNVLCGEREEARACHATHNRACRCRTGFFAHAG-----FLEHASCPPGAGVIA 57
DB 57 CLSCGSSCDVETQACTREONRICTCPGWYCALSKOEGCRICAPLKRCPGFGVAR 116
QY 58 GTPSONTOCOPCPGPTGTFSSASSSSSEOCOPHRNCTALGIALNVPSSSHDTLCSTGFP 117
DB 117 GRETSDVCKPCAPGPTFTNTSTSDICRPHQCNVNA-----IPGNASMDAVCTST--PT 170
QY 118 STRVGA 124
DB 171 RSMARGA 177

```

```

RESULT 7
Q62327 PRELIMINARY: PRT: 459 AA.
AC 062327:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NOD;
RX MEDLINE: 95178848.
RA POWELL E.E., WICKER L.S., PETERSON L.B., TODD J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
gene.";
RL MAMM. GENOME 5:726-727(1994).
DR EMBL: X76401; G433831; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PFAM: PF00020; TNFR_c6; 4.
FT NON-TER 1
FT VARIANT 87 1 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 57791809 CRC32;

```

```

Query Match 26.3%; Score 221.5; DB 10; Length 459;
Best Local Similarity 37.2%; Pred. No. 1.5e-09;
Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

```

```

QY 4 CRYCNVLCGEREEARACHATHNRACRCRTGFF---AHAGF---CLEHASCPPGAGVIA 56
DB 79 CLSCGSSCDVETQACTREONRICTCPGWYCALSKOEGCRICAPLKRCPGFGVAR 138
QY 57 PETPSONTNOCOPCPGPTGTFSSASSSSSEOCOPHRNCTALGIALNVPSSSHDTLC 109
DB 139 SRAPGNVLCRACAPGTSTSTVDCRPHRICSLA---IPGNASMDAVC 187

```

```

RESULT 8
Q88734 PRELIMINARY: PRT: 482 AA.
AC 088734:
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS MUS MUSCULUS (MOUSE).

```

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA HURLE B., SEGAGE F., RODRIGUEZ R., RAMOS S.S., LAZO P.S.;
 RT "The Mouse Tumor Necrosis Factor Receptor Gene: genomic Structure and
 RT Characterization of the two transcripts."
 RL GENOMICS 0:0-0(0).
 DR EMBL: Y14619; E1299292; -.
 DR EMBL: Y14620; E1299292; JOINED.
 DR EMBL: Y14621; E1299292; JOINED.
 DR EMBL: Y14622; E1299292; JOINED.
 DR EMBL: Y14623; E1299292; JOINED.
 DR EMBL: Y14679; E1299292; JOINED.
 DR PROSITE: PS00652; TNFR_NGFR.1; 2.
 SQ SEQUENCE 482 AA; 5106 MW; 3A3069B5 CRC32;

Query Match 26.3%; Score 221.5; DB 10; Length 482;
 Best Local Similarity 37.2%; Pred. No. 1.6e-09;
 Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

OY 4 CIRCUNVLGGEREERARACHATHNRACRCRTGFF---AHAGF---CLEHASCPPGAGVIA 56
 DB 101 CLSCSSCSTDOYETRACTCOQNRVCCEAGRYCALTHSGSCRCQMRISKCGPFGVAS 160
 OY 57 PCPTSONTOCOPCPGPFSSASSSSSEOCOPHRNCTALGLALNVGSSSHDTLCISCTG 109
 DB 161 SRAPNGNVLKACAPGTFSDTSTDCVCRPHRICISLA---IPGNASTDAVC 209

RESULT 9
 ID 075509 PRELIMINARY; PRT; 655 AA.
 AC 075509;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE TNFR-RELATED DEATH RECEPTOR-6.
 GN DR6.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA PAN G., BAUER J.H., HARIDAS V., WANG S., LIU D., NI J., YU G.,
 RA VINCENTZ C., AGGARWAL B.B., DIXIT V.M.;
 RT "Identification and functional characterization of DR6, a novel death
 RT domain-containing TNF receptor."
 RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF068868; G3549263; -.
 SQ SEQUENCE 655 AA; 71844 MW; 67773C46 CRC32;

Query Match 22.8%; Score 192; DB 2; Length 655;
 Best Local Similarity 32.8%; Pred. No. 3.2e-07;
 Matches 39; Conservative 18; Mismatches 62; Indels 0; Gaps 0;

OY 1 LERCRCNVLCGEREERARACHATHNRACRCRTGFF-----AHAGFCLHASCPPGAGV 60
 DB 103 IECHCHCSOPCPMPMEIKLPCALITDRECTCPGMEFSNATCAPHTVCGVGRKKGTE 162
 OY 61 SONTOCOPCPGPFSSASSSSSEOCOPHRNCTALGLALNVGSSSHDTLCISCTGFP 119
 DB 163 TEDYRKQCKARGTFSVPSSVMCKATDCLSONLVYIKGTETDNCGLTLPSSSST 221

RESULT 10
 ID 057277 PRELIMINARY; PRT; 348 AA.
 AC 057277;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)

DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1996(96-17), ZAIRE-1996 (96-16);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88543; G2738105; -.
 DR EMBL: U87841; G2738061; -.
 DR PROSITE: PS00652; TNFR_NGFR.1; 2.
 SQ SEQUENCE 348 AA; 38212 MW; 1AAAB869 CRC32;

Query Match 22.4%; Score 188.5; DB 11; Length 348;
 Best Local Similarity 31.3%; Pred. No. 3.3e-07;
 Matches 41; Conservative 24; Mismatches 59; Indels 7; Gaps 2;

OY 1 LERCRCNVLCGEREERARACHATHNRACRCRTGFF-----AHAGFCLHASCPPGAGV 54
 DB 80 LQACLSNCRCDSONOYETRSCNTHNRICEGSPGYCLLKSGSCRTCTSKTKCGIGYGV 139
 OY 55 IAPGTPSONTOCOPCPGPFSSASSSSSEOCOPHRNCTALGLALNVGSSSHDTLCISCTG 114
 DB 140 -SGYSTGTVICSPCGPGYSHTVSSTDKCEPVTNTFNYIDVEINLYPVNDTSCRITTT 198
 OY 115 FPLSTRVPGAE 125
 DB 199 TGLSEISTSE 209

RESULT 11
 ID 057103 PRELIMINARY; PRT; 348 AA.
 AC 057103;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1979 (79-0005);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87847; G2738073; -.
 DR PROSITE: PS00652; TNFR_NGFR.1; 2.
 SQ SEQUENCE 348 AA; 38184 MW; EF87B2D2 CRC32;

Query Match 22.4%; Score 188.5; DB 11; Length 348;
 Best Local Similarity 31.3%; Pred. No. 3.3e-07;
 Matches 41; Conservative 24; Mismatches 59; Indels 7; Gaps 2;

OY 1 LERCRCNVLCGEREERARACHATHNRACRCRTGFF-----AHAGFCLHASCPPGAGV 54
 DB 80 LQACLSNCRCDSONOYETRSCNTHNRICEGSPGYCLLKSGSCRTCTSKTKCGIGYGV 139
 OY 55 IAPGTPSONTOCOPCPGPFSSASSSSSEOCOPHRNCTALGLALNVGSSSHDTLCISCTG 114
 DB 140 -SGYSTGTVICSPCGPGYSHTVSSTDKCEPVTNTFNYIDVEINLYPVNDTSCRITTT 198
 OY 115 FPLSTRVPGAE 125
 DB 199 TGLSEISTSE 209

```

Db      198 TGLSEISTSE 208

RESULT 14
ID      057100      PRELIMINARY;      PRT;      349 AA.
057100
AC      057100.
DT      01-JUN-1998 (TREMBLREL. 06, CREATED)
DT      01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT      01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE      TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN
NM
OS      MONKEYPOX VIRUS.
OC      VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRIDAE; CHORODOPOXYIRINAE;
OC      ORTHOPOXYVIRUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-NIGERIA-1971 (71-0082).
RA      LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
      SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
      EMBL; U87844; G2738067; -.
DR      PROSITE; PS00652; TNFR_NGFR_1; 2.
SQ      SEQUENCE      349 AA; 38239 MW; F8871DD2 CRC32;

Query Match      21.6%; Score 184; DB 11; Length 349;
Best Local Similarity 30.6%; Pred. No. 7e-07;
Matches 41; Conservative 24; Mismatches 57; Indels 12; Gaps 4;

QY      1 LERCYCNVLGGEREEERACHATINRACRCHTGF-----AHNGFLEHASCPPGAGV 54
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      80 LQACSCNGCRCDNSOVETRSCNTTHNRICECSPGYCLLKAGSGCSTCISKTKGIGYGV 139
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      55 IAPGPTSGTQCCPPPTFFSASSSSQCCOP---HRNCTALGLALNYPGSSSHDTLCTS 111
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      140 -SGYISTGVDVICSPPGPGFTSHYTSYDCKEPEVYNSFNFIYDVEINL--YVNDTSCIR 196
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      112 CTGPELSTRVPGAE 125
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      197 TTTTGLSEISTSE 210

RESULT 15
ID      057291      PRELIMINARY;      PRT;      349 AA.
057291
AC      057291.
DT      01-JUN-1998 (TREMBLREL. 06, CREATED)
DT      01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT      01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE      TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN
NM
OS      MONKEYPOX VIRUS.
OC      VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRIDAE; CHORODOPOXYIRINAE;
OC      ORTHOPOXYVIRUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-VARIOUS STRAINS;
RA      LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
      SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
      EMBL; U88144; G2738087; -.
DR      EMBL; U87842; G2738063; -.
DR      EMBL; U87994; G2738073; -.
DR      EMBL; U87995; G2738081; -.
DR      EMBL; U88143; G2738085; -.
DR      PROSITE; PS00652; TNFR_NGFR_1; 2.
SQ      SEQUENCE      349 AA; 38295 MW; 7313FCF9 CRC32;

Query Match      21.6%; Score 182; DB 11; Length 349;
Best Local Similarity 30.6%; Pred. No. 9.9e-07;
Matches 41; Conservative 24; Mismatches 57; Indels 12; Gaps 4;

```

```

OY 1 LERCRYCNVLCGEREEEARACHATHNRACRRTGFF-----AHAGFCLERHASCPPGAGV 54
    : | | | : | | | | | | : | | |
Db 80 LQACLSGNGRCDSDNQVETRSCNTHNRICECSPGYICLLKGASGCRCTCISKTKCGIGYV 139
    : | | | | | | | | | | : | | |
OY 55 IAPGPPSONTOQCPGPPGTFSSASSSSSECCOP---HRNCTALGLALNVPGSSSHDTLCTS 111
    : | | | | | | | | | | : | | |
Db 140 -SGYTSIGDVICSPGPGTYSHVSSSTDCKECPVTSNFTFNVIDEINT--YFVNDTSCTR 196
    : | | | | | | | | | | : | | |
OY 112 CTGFPLSTRVPGAE 125
    : | | | | | | | | | | : | | |
Db 197 TTTTGLSESISTSE 210
    : | | | | | | | | | | : | | |

```

Search completed: May 14, 1999, 21:39:57
 Job time: 7991 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 17:21:17 ; Search time 42.79 Seconds
(without alignments)
99.260 Million cell updates/sec

Title: US-09-212-270-3
Perfect score: 1162
Sequence: 1 MALKVLP.LHRTVLF.AAILFL.....STANGSSWRNRP.VAVSHPSV 210

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database: A_Geneseq_34.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	869	74.8	417	1	W70386 Amino acid sequenc
2	869	74.8	423	1	W70387 Amino acid sequenc
3	155	13.3	438	1	R81882 Plasmid pDC406/OX4
4	155	13.3	206	1	R81881 Mouse type-II memb
5	155	13.3	438	1	W48976 OX40/Tc muclein. Pu
6	155	13.3	206	1	W48977 Mouse OX40 extrace
7	151	13.0	176	1	W80254 Amino acid sequenc
8	142.5	12.3	283	1	W05809 Human tumour necro
9	142.5	12.3	283	1	W69238 Hepesvirus entry
10	142	12.2	625	1	W69238 Murine NF-kB recep
11	142	12.2	625	1	W68294 Human herpes simpli
12	141.5	12.2	283	1	W12659 Compox virus T2-eq
13	141	12.1	355	1	R85072 Myxoma virus T2 pr
14	140	12.0	326	1	R85072 Human TNF receptor
15	139.5	12.0	283	1	W60045 A receptor protein
16	135	11.6	616	1	W79233 NF-kB receptor act
17	133	11.4	451	1	W69952 RANK polypeptide p
18	133	11.4	451	1	W69951 NF-kB receptor act
19	133	11.4	451	1	W68287 RANK polypeptide p
20	133	11.4	451	1	W68288 NF-kB receptor act
21	130.5	11.2	836	1	W64486 Human DR3 protein.
22	127.5	11.0	197	1	W23220 Extracellular doma
23	126.5	10.9	466	1	R07447 Human laminin B1 c
24	126.5	10.9	186	1	R62655 Compox virus Pst I
25	126.5	10.8	461	1	R07450 Rat Tumour Necrosi
26	125.5	10.8	1786	1	W50893 Human laminin B1 c
27	124.5	10.7	277	1	R38859 CD40 protein.
28	124.5	10.5	417	1	W31517 Death domain conta
29	121.5	10.5	428	1	W31516 Human apoptosis pr
30	121.5	10.5	417	1	W26709 Human apoptosis pr
31	121.5	10.5	417	1	W57045 Human apoptosis in
32	121.5	10.5	417	1	W57045 Human apoptosis in
33	121	10.4	181	1	W26708 Mouse apoptosis in
34	120.5	10.4	411	1	W57046 Human apoptosis in
35	120	10.3	285	1	W33359 TBK(20-190)/hCG-a1
36	119.5	10.3	1764	1	P91672 Primary amino acid
37	119.5	10.3	518	1	R51003 Sequence of a reco
38	119.5	10.3	1766	1	R51004 p75 Tumour Necrosi
39	119.5	10.3	1766	1	W50894 Mouse laminin B1 c
40	118.5	10.2	461	1	R11141 Human TNF-R deduce
41	118.5	10.2	485	1	R24016 Fusion protein TNF
42	118.5	10.2	461	1	R42058 Fliorblast derived
43	117.5	10.1	256	1	W33357 TBP(20-161)/hCG-a1

44 117 10.1 277 1 R76996
45 117 10.1 1798 1 W50896

Deduced sequence e
Human laminin B2 c

ALIGNMENTS

RESULT 1	
ID W70386	standard; Protein; 417 AA.
AC W70386	
DE 02-DEC-1998 (first entry)	
DI Amino acid sequence of human alpha-OAF065.	
KW Human; alpha-OAF065; stroma cell; antibody; inflammatory;	
KW cytokine-mediated disease; rheumatism; ulcerative colitis.	
OS Homo sapiens.	
FH Key	Location/Qualifiers
FT Misc_difference 223	/note= "encoded by AGA"
FT Misc_difference 224	/note= "encoded by CCT"
FT Misc_difference 224	/note= "encoded by CCT"
PN W09838304-A1.	
PD 03-SEP-1998.	
PE 26-FEB-1998; J00799.	
PF 27-FEB-1997; JP-043143.	
PR (ONOX) ONO PHARM CO LTD.	
PI Fukushima D, Konishi M, Tada H;	
DR WPI: 98-481205/41.	
DR N-PSDB: V33361.	
PT Membrane polypeptide expressed by human stroma cells, and antibodies	
PT recognising it - for treatment of inflammatory and other	
PT cytokine-mediated diseases.	
PS Claim 1; Pages 28-30; 54pp; Japanese.	
CC This is the amino acid sequence of the human alpha-OAF065, used in	
CC the method of the invention. The process involves the use of peptides	
CC expressed by stroma cells, and its antibodies are used for in the	
CC prevention and treatment of inflammatory and other cytokine-mediated	
CC diseases such as rheumatism, ulcerative colitis.	
CC Sequence 417 AA;	
SQ	
Query Match	74.8%; Score 869; DB 1; Length 417;
Best Local Similarity	83.7%; Pred. No. 6.2e-70;
Matches 154; Conservative 11; Mismatches 19; Indels 0; Gaps 0;	
QY 1 MALKVLP.LHRTVLF.AAILFL.LACKVSCETGDCRQEFKDRSGNCVLC.KOCGPGWELSK 60	
DB 1 MALKVLP.LHRTVLF.AAILFL.LACKVSCETGDCRQEFKDRSGNCVLC.KOCGPGWELSK 60	
QY 61 ECGFGYGEDAQCVPGRPHRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120	
DB 61 ECGFGYGEDAQCVPGRPHRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120	
QY 61 ECGFGYGEDAQCVPGRPHRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120	
DB 61 ECGFGYGEDAQCVPGRPHRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120	
QY 121 FFRKTKLVGFQDMCEVPCGDDPPPEPHCTSKVNLKYSSTVSSPRDRAVAATGALAT 180	
DB 121 FFRKTKLVGFQDMCEVPCGDDPPPEPHCTSKVNLKYSSTVSSPRDRAVAATGALAT 180	
QY 181 VLLA 184	
DB 181 VLLA 184	
RESULT 2	
ID W70387	standard; Protein; 423 AA.
AC W70387	
DE 02-DEC-1998 (first entry)	
DI Amino acid sequence of human beta-OAF065.	
KW Human; beta-OAF065; stroma cell; antibody; inflammatory;	
KW cytokine-mediated disease; rheumatism; ulcerative colitis.	
OS Homo sapiens.	
FH Key	Location/Qualifiers
FT Misc_difference 223	/note= "encoded by AGA"

FT Misc difference /note= "encoded by AGA"
FT 224 /note= "encoded by CCR"
PN MO9838304-A1.
PD 03-SEP-1998.
PF 26-FEB-1998; J00799.
PR 27-FEB-1997; JP-043143.
PA (ONOF) ONO PHARM CO LTD.
PI Fukushima D, Konishi M, Tada H;
DR WPI: 98-481205/41.
N-PSDB: V33362.
PT Membrane polypeptide expressed by human stroma cells, and antibodies
PT recognizing it - for treatment of inflammatory and other
PT cytokine-mediated diseases.
PS Disclosure: Pages 37-49; 54pp; Japanese.
CC This is the amino acid sequence of the human beta-OAF065, used in
CC the method of the invention. The process involves the use of peptides
CC expressed by stroma cells, and its antibodies are used for in the
CC prevention and treatment of inflammatory and other cytokine-mediated
CC diseases such as rheumatism, ulcerative colitis.
SQ Sequence 423 AA;

Query Match 74.8%; Score 869; DB 1; Length 423;
Best Local Similarity 83.7%; Pred. No. 6.3e-70;
Matches 154; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 1 MALKVLPLHRTVLFALILLLHACKVSCETGDCRQEFKDRSGNCVLCQGPMELSK 60
ID 1 MALKVLLEDEKFFFLVLLGVISCKVCTETGCRQEFKDRSGNCVPCNOCQGPMELSK 60
DB 1 MALKVLLEDEKFFFLVLLGVISCKVCTETGCRQEFKDRSGNCVPCNOCQGPMELSK 60
OY 61 ECGFGYGDADQVCPRHFRKEDMGFQKCKPCADALVNRFORANCSSHTSDAVCGDCLPG 120
ID 61 ECGFGYGDADQVCTGLHFKEDMGFQKCKPCADALVNRFORANCSSHTSDAVCGDCLPG 120
DB 61 ECGFGYGDADQVCTGLHFKEDMGFQKCKPCADALVNRFORANCSSHTSDAVCGDCLPG 120
OY 121 FYRKTKLVGFQDMCEVPCGDPPEPYRPHCTSKYLVKISSVSSPRDTAAVAAYIC 180
ID 121 FYRKTKLVGFQDMCEVPCGDPPEPYRPHCTSKYLVKISSVSSPRDTAAVAAYIC 180
DB 121 FYRKTKLVGFQDMCEVPCGDPPEPYRPHCTSKYLVKISSVSSPRDTAAVAAYIC 180
OY 181 VLLA 184
ID 181 VLLA 184
DB 181 VLLA 184

RESULT 3
R81882
ID R81882 standard; Protein: 438 AA.
AC R81882;
DE 30-MAR-1996 (first entry)
DE Plasmid pDC406/OX40/Fc* encoding an OX40/Fc mutein protein.
KW OX40; OX40-L; cytokine; cell surface molecule; plasmid.
KW pDC406/OX40/Fc*; membrane glycoprotein.
OS Synthetic.
PN US5457035-A.
PD 10-OCT-1995.
PF 23-JUL-1993; 097827.
PR 23-JUL-1993; US-097827.
PA (IMMV) IMMUNEX CORP.
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR WPI: 95-357992/46.
N-PSDB: T00829.

PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
PS Example 2: Column 35-38; 26pp; English.
CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,
CC and is used to express a soluble OX40/Fc mutein fusion protein for
CC use in detecting CDNA clones encoding a OX40 ligand. The Fc
CC fragment may be derived from human IgG1, and the plasmid may be
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
CC line. Culture supernatant was purified by affinity chromatography
CC and this was used, together with labeled goat anti-human IgG to
CC screen various cell lines.

SQ Sequence 438 AA;

Query Match 13.3%; Score 155; DB 1; Length 438;
Best Local Similarity 27.6%; Pred. No. 1.9e-06;
Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

OY 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
ID 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
DB 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
OY 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
ID 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
DB 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
OY 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
ID 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
DB 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
OY 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
ID 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
DB 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
OY 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
ID 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
DB 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
OY 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
ID 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
DB 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
OY 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205
ID 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205
DB 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205

RESULT 4
R81881
ID R81881 standard; Protein: 206 AA.
AC R81881;
DE 08-JUL-1996 (first entry)
DE Mouse type-II membrane polypeptide OX40 extracellular domain.
KW OX40; OX40-L; cytokine; cell surface molecule;
KW membrane glycoprotein.
OS Mus musculus.
PN US5457035-A.
PD 10-OCT-1995.
PF 23-JUL-1993; 097827.
PR 23-JUL-1993; US-097827.
PA (IMMV) IMMUNEX CORP.
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR WPI: 95-357992/46.
N-PSDB: T00826.

PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
PS Example 1: Column 33-34; 26pp; English.
CC This sequence encodes the extracellular domain of OX40, a membrane
CC glycoprotein present on the CD4 positive subset of activated T
CC cells.
SQ Sequence 206 AA;

Query Match 13.3%; Score 155; DB 1; Length 206;
Best Local Similarity 27.6%; Pred. No. 8.1e-07;
Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

OY 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
ID 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
DB 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMELSKEGFGYGDADQVPCR 76
OY 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
ID 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
DB 77 PHRFKEDMGFQKCKPCADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD--- 133
OY 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
ID 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
DB 65 TGFYNANVAYDICKQCTOCNHNHSGSEIKONCPTDQTYC-RCRPQTQPR-----QDSGY 117
OY 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
ID 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
DB 133 ---MECVPCGDPPEPYR---HCTSKYLVKISSVSSPRDTAAVAAYIC---SALATVL 182
OY 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
ID 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
DB 118 KLGVDQVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
OY 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
ID 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
DB 183 LACS-----SCVSSYARSSSRKRRPAYSHPVS 210
OY 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205
ID 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205
DB 175 WETQRPFRPTVQST--TWMPRTSELPSTPL 205

RESULT 5

Page 7

```

CC to its receptor. T2 protein is useful (when administered by
CC injection or infusion from sustained release implants, etc.) for
CC treating TNF associated toxicity e.g. side effects of using TNF as
CC an antitumour agent or in cases of cachexia and septic shock where
CC TNF production is excessive.
SQ Sequence 326 AA;

Query Match 12.0%; Score 140; DB 1; Length 326;
Best Local Similarity 25.0%; Pred. No. 2.9e-05;
Matches 51; Conservative 20; Mismatches 81; Indels 52; Gaps 12;

QY 13 LEAALLFLHLAC-----KVCSETGDCRQGEFFKDRSSNCVLCCKCGGMEISKCCGCGY 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MERLLILLAYVACVGGAGAPYAGDRGKCRGNDY -EKDILC--CTSCPPGYSASRLC--GP 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 67 GBDACVCCPRPRHREEDNGF-QKCRPCADCALVYNFQANCSHTSDAIVCGDCLPFGY--- 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 GSDIVCSPCKNKETFPASTINHPACVSCGRGRTGHTISEQSDKRTDRVC-DCSAGNYCLL 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 123 -----RRTKL-VGF-----QDMECVPCGDGPPPY-----DPHCTSKYNL 155
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 KGOECRICAPKTKCPAGICVSGHTRTDVLCCTK---PRITYSDAVSTETCTSFENY 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 156 VKISSTVSSPRTAAVAIVCSALA 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 13, 1999, 17:21.18
Job time: 20717 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:04 ; Search time 33.52 Seconds

(without alignments)
234.686 Million cell updates/sec

Title: US-09-212-270-3

Sequence: 1 MALNVLPHRIVLFAILFL.....STARSSWRNPVAVSHPSV 210

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database: PIR_58:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161	13.9	271	2	SI2783
2	155	13.3	272	2	148700
3	140.5	12.1	435	2	I54182
4	140	12.0	326	1	GOVZML
5	138.5	11.9	454	1	GQMS71
6	138.5	11.9	454	2	I57826
7	138.5	11.9	349	2	D36858
8	132	11.4	416	2	JN0006
9	129	11.1	461	1	GQRT71
10	128.5	11.1	277	2	A60771
11	125.5	10.8	250	1	A49053
12	125.5	10.8	1786	1	MMHUB1
13	120.5	10.4	425	2	A26431
14	119.5	10.3	1786	1	MMMSB1
15	118.5	10.2	427	1	GQHDN
16	118.5	10.2	461	2	A35356
17	117	10.1	260	1	A46517
18	116	10.0	277	2	I37552
19	115.5	9.9	334	2	A48151
20	114.5	9.9	869	1	JC4858
21	114.5	9.9	1801	1	MMRTS
22	114	9.8	2482	2	A47648
23	114	9.8	2277	2	B47648
24	114	9.8	289	2	A46515
25	114	9.8	305	2	A46476
26	110.5	9.5	873	1	ORRBVD
27	110.5	9.5	837	2	S43656
28	110	9.5	1798	2	S53869
29	109.5	9.4	873	1	A49729
30	109.5	9.4	863	1	S51789
31	109.5	9.4	1607	1	MMMSB2
32	109	9.4	596	2	A45664
33	108.5	9.3	873	1	I48952
34	108	9.3	1790	1	MMFEB1
35	108	9.3	1797	2	A55677
36	108	9.3	255	2	I38426
37	108	9.3	255	2	J70752
38	107.5	9.3	3084	1	MMMSA
39	107.5	9.3	667	2	A48579

40	107.5	9.3	1680	2	A43434	furin (EC 3.4.21.7
41	107.5	9.3	265	2	B48151	sperm tail protein
42	107	9.2	1548	2	S34583	serine proteinase
43	106.5	9.2	461	2	JC4302	tumor necrosis fac
44	105.5	9.1	3712	2	S18253	laminin alpha-1 ch
45	104.5	9.0	1609	1	MMHUB2	laminin gamma-1 ch

ALIGNMENTS

RESULT 1

SI2783

OX40 antigen precursor - rat

N:Alternate names: nerve growth factor receptor homolog

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997

C:Accession: SI2783; S08036

R:Maillet, S.; Fossum, S.; Barclay, A.N.

EMBO J. 9, 1063-1068, 1990

A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphoc

A:Reference number: SI2783; MUID:90214614

A:Accession: SI2783

A:Molecule type: mRNA

A:Residues: 1-271 <MAL>

A:Cross-references: EMBL:X17037; NID:957830; PID:957831

C:Keywords: growth factor receptor; transmembrane protein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-271/Product: OX40 antigen #status predicted <MAT>

F:211-235/Domain: transmembrane #status predicted <TM>

Query Match 13.9%; Score 161; DB 2; Length 271;

Best Local Similarity 27.2%; Pred. No. 1.2e-06;

Matches 53; Conservative 26; Mismatches 72; Indels 44; Gaps 11;

QY	45	NCV-----LCKOCGGMELSKCEGFGEDACVPCPRHREKEDMGFOKPCADC	95
DB	25	NCVKDTPSGHKCCRCQCPGHGVSRCC--DHTRDTCVCHPCCEGFYNEAVNYDCKOCTOC	82
QY	96	ALVNRQ-RANGSHSDAVCGDCLPGFRTKLVGQDMECVPCGDPDPPEP-----	148
DB	83	NHRSGELKQNTPTEDTVC-QCRPGTORODSSKRLGVDCVPC--PRGHFSGSNQACK	139
QY	148	---HCTSKYLVKISSTVSSPRDTAAVVIC---SALATVLLACS-----SCVSTARG	195
DB	140	PMTNCR--LSGQIRHPANSNDT-----VCEDRSILATLIMETQRTTRPTTVPST----	190
QY	196	SSWRNPVAVSHPSV 210	
DB	190	TWPFRTSQLPSTPTL 204	

RESULT 2

148700

gene ox40 protein - mouse

N:Alternate names: OX40 antigen

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996

C:Accession: I48700; I48334; S34377

R:Calderhead, D.M.; Buhlmann, J.E.; van den Bertwegh, A.J.; Claassen, E.; Noelle, R.J.

J. Immunol. 151, 5261-5271, 1993

A:Title: Cloning of mouse OX40: a T cell activation marker that may mediate T-B cell

A:Reference number: I48700; MUID:94044750

A:Accession: I48700

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-272 <RES>

A:Cross-references: EMBL:Z21674; NID:9312827; PID:9312828

R:Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.

Eur. J. Immunol. 25, 926-930, 1995

A:Title: Gene structure and chromosome localization of the mouse homologue of rat OX

A:Reference number: I48334; MUID:95255413

R:Rothe, J.G.: Brockhaus, M.; Gentz, R.; Lesslauer, W.
 Immunogenetics 34, 338-340, 1991
 A:Title: Molecular cloning and expression of the mouse tnfr receptor type b.
 A:Reference number: S19021; MUID:92039815
 A:Accession: S19021
 A:Molecule type: mRNA
 A:Residues: 1-454 <ROT>
 A:Cross-references: EMBL:X57796; NID:g54848; PID:g54849
 R:Bebo, B.F.
 Immunogenetics 39, 450-451, 1994
 A:Title: Nucleotide sequence of the TNF type I receptor from a mouse endothelioma cell 1
 A:Reference number: 154532; MUID:94245292
 A:Accession: 154532
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-454 <RES>
 A:Cross-references: GB:L26349; NID:g430732; PID:g430733
 C:Comment: This protein is one of two distantly related receptors for both TNF-alpha (ca
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-454/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-212/Domain: extracellular #status predicted <EXT>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG3>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:213-235/Domain: transmembrane #status predicted <MEM>
 F:236-454/Domain: intracellular #status predicted <INT>
 F:54,151,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.9%; Score 138.5; DB 1; Length 454;
 Best Local Similarity 22.0%; Pred. No. 0.00013;
 Matches 56; Conservative 31; Mismatches 103; Indels 65; Gaps 11;

```

QY 6 LPLRYLFAIFLFLHLACKVSCETG-----DCROEFKDSGNCVLCQK 52
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 3 LPTVPGLLSLVLLALMGHPGSGVTLVPSLDREKRDSCPGKTVHSHKNNISCTK 62

QY 53 GPGMELSKECGFYGEDACVPCRPFRKEDMGF-QKCKPCADAL-VNRFORANCST 110
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 63 HKGYLVSDCP-SPGRTVCRECKGFTASQNVLRQCLCKCRKMSQVEISPCQADK 121

QY 111 DAVCG-----DCLPGFYRKTKL-----VGF--QDMCCVP 137
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 122 DTVCCKENOFRLSETHFCVDCSPCFNGTVPICKETQNTVCNCHAGFLESECV 181

QY 138 CGDPPPEPEHCTSKVNLVKS-----STVSSPDITVAIVGALATVLLACSSC-VSS 191
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 182 CS-----HCKKNECKMLCLPPPLAVNTNPDSGTAVLLPLVILGLCLSFIFISL 233

QY 192 TARGSSWRNPVAVSS 206
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 234 MCRTPRMR--PEVYS 246
```

RESULT 6
 157826
 tumor necrosis factor receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-May-1997
 C:Accession: 157826
 R:Rothe, J.G.; Bluetmann, H.; Gentz, R.; Lesslauer, W.; Steinmetz, M.
 Mol. Immunol. 30, 165-176, 1993
 A:Title: Genomic organization and promoter function of the murine tumor necrosis factor
 A:Reference number: 157826; MUID:93156721
 A:Accession: 157826
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-454 <RES>
 A:Cross-references: GB:M76656; NID:g202100; PID:g202102
 C:Genetics:

A:Gene: TNFR-2
 A:Intons: 13/3; 65/1; 108/1; 158/1; 184/2; 210/1; 248/1; 257/3; 353/1
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: cytokine receptor

Query Match 11.9%; Score 138.5; DB 2; Length 454;
 Best Local Similarity 22.0%; Pred. No. 0.00013;
 Matches 56; Conservative 31; Mismatches 103; Indels 65; Gaps 11;

```

QY 6 LPLRYLFAIFLFLHLACKVSCETG-----DCROEFKDSGNCVLCQK 52
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 3 LPTVPGLLSLVLLALMGHPGSGVTLVPSLDREKRDSCPGKTVHSHKNNISCTK 62

QY 53 GPGMELSKECGFYGEDACVPCRPFRKEDMGF-QKCKPCADAL-VNRFORANCST 110
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 63 HKGYLVSDCP-SPGRTVCRECKGFTASQNVLRQCLCKCRKMSQVEISPCQADK 121

QY 111 DAVCG-----DCLPGFYRKTKL-----VGF--QDMCCVP 137
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 122 DTVCCKENOFRLSETHFCVDCSPCFNGTVPICKETQNTVCNCHAGFLESECV 181

QY 138 CGDPPPEPEHCTSKVNLVKS-----STVSSPDITVAIVGALATVLLACSSC-VSS 191
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 182 CS-----HCKKNECKMLCLPPPLAVNTNPDSGTAVLLPLVILGLCLSFIFISL 233

QY 192 TARGSSWRNPVAVSS 206
   || : : : : : || : : : : : || : : : : : || : : : : :
Db 234 MCRTPRMR--PEVYS 246
```

RESULT 7
 D36858
 gene G4R protein - variola virus
 N:Alternate names: B28R protein (COP)
 C:Species: variola virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
 C:Accession: D36858; S46888; S32385; S35987
 R:Blinov, V.M.
 submitted to Genbank, November 1992
 A:Description: not shown.
 A:Reference number: D36858
 A:Accession: D36858
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <BLI>
 A:Cross-references: GB:X69198; NID:g456758; PID:g4567087
 A:Experimental source: strain India-1967, ssp. major, isolate Ind3
 R:Kolymhalov, A.A.; Blinov, V.M.; Gytorov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
 submitted to the EMBL Data Library, April 1992
 A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
 A:Reference number: S46888
 A:Accession: S46888
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <KOL>
 A:Cross-references: EMBL:X67117; NID:g516428; PID:g516449
 A:Experimental source: strain India-1967, isolate Ind3
 R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
 FEBS Lett. 319, 80-83, 1993
 A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protect
 A:Reference number: S32385
 A:Accession: S32385
 A:Molecule type: DNA
 A:Residues: 31-168 <SHC>
 A:Cross-references: EMBL:X69198
 A:Experimental source: strain India-1967, ssp. major
 C:Genetics:
 A:Gene: G4R
 C:Superfamily: NGF receptor repeat homology
 F:68-109/Domain: NGF receptor repeat homology <NG2>
 F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match 11.9%; Score 138.5; DB 2; Length 349;
 Best Local Similarity 23.7%; Pred. No. 0.00011;
 Matches 50; Conservative 30; Mismatches 88; Indels 43; Gaps 10;

Oy 10 RTVFAALLFLIHL-----ACKVSCETGDCROQEFKDRSGNCVLCCKGCGMELSKREG 63
 Db 2 KSVLLYLILFLSCIIINGRDAPYTPPPKCKDTEYK-RHNLIC--CLSCPGETVASRLC- 58
 Oy 64 FGFGDACCPCPRPHRF-KEDMGFQCKRCPADCALVNRFORANCSTSDAVGDCDLPFGY 122
 Db 58 -DSKTNCTCTGCGSGTFSRNHNLPACLSCNGRCNSNVETRSCTTHNRIC-ECSPGY 115
 Oy 123 -----RRTKL-VGF-----QMECVPCG-----DPPPEPHCTSK 152
 Db 116 CLKSSGCKACVSGTKGIGVSGHVSIGVTSFGFGYTSHTVSADKCEVPVNNNT 175
 Oy 153 VNLVKISSTVSPRDTAAVAVICALATVLL 183
 Db 176 FNYIDVEITLVPNDTCTRTTTLGISEIL 206

RESULT 8

UN0006
 nerve growth factor receptor, low affinity - chicken
 N:Alternate names: NGF receptor
 C:Species: Gallus gallus (chicken)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 01-May-1998
 C:Accession: JN0006; A60504
 R:Large, T.H.; Weiskamp, G.; Heider, J.C.; Radeke, M.J.; Misko, T.P.; Shooter, E.M.; Reid
 Neuron 2, 1123-1134, 1989
 A:Title: Structure and developmental expression of the nerve growth factor receptor in t
 A:Reference number: JN0006; MUID:90166579
 A:Accession: JN0006
 A:Molecule type: mRNA
 A:Residues: 1-416 <L&R>
 A:Experimental source: embryonic chick brain
 R:Heuer, J.G.; Fawcett-Najle, S.; Wheeler, E.F.; Bothwell, M.
 Dev. Biol. 137, 287-304, 1990
 A:Title: Structure and developmental expression of the chicken NGF receptor.
 A:Reference number: A60504
 A:Accession: A60504

A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 21-35,'Y','37-172,'K','174-275,'S','277-395,'R','397-416 <HEU>
 C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
 C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
 C:Comment: This protein is thought to form a high-affinity receptor when it associates w
 C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor;
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-416/Product: nerve growth factor receptor #status predicted <MAT>
 F:21-239/Domain: extracellular #status predicted <EXT>
 F:24-57/Domain: NGF receptor repeat homology <NG1>
 F:59-100/Domain: NGF receptor repeat homology <NG2>
 F:101-139/Domain: NGF receptor repeat homology <NG3>
 F:141-181/Domain: NGF receptor repeat homology <NG4>
 F:189-237/Region: serine/threonine-rich
 F:240-261/Domain: transmembrane #status predicted <MEM>
 F:262-416/Domain: intracellular #status predicted <INT>
 F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.4%; Score 132; DB 2; Length 416;
 Best Local Similarity 27.8%; Pred. No. 0.00043;
 Matches 59; Conservative 19; Mismatches 100; Indels 34; Gaps 11;

Oy 21 LHLACKVSCETGD---CROGEFKDR-SGNCVLCCKGCGMELSKREGCGY-----GPD 69
 Db 84 LHSMAPCVESPDAVCAKAYGFQDELSSGCKECSIC-----EVGGLMFPDRSDSD 135
 Oy 70 AOCVPCPRHREKDWG-OKCKPCADCALVNRFORANCSTSDAVGDCDLPFGYKRT-KL 127

Db 136 TVCECPPEGTSDANFVDPCLPCTICE-ENEMVWKECTADSAEBCRDRLHPRTHTHPSL 194
 Oy 128 VGFQDMEVPCGDDPPPEPHCTSKVNLVKSIVSSP---RDTA--VAAVICALATVL 182
 Db 195 AGSDSP--PITRDFENEGMATTADIVTVMGSSQPVSRGTADNLIPYCSILAAYV 252
 Oy 183 LACSSCV-----SSRANGSSWRNPVASSHPS 209
 Db 253 VGLVAVIAFKRMNSCKOKNGCANRPVNOTPS 284

RESULT 9

GORT1
 tumor necrosis factor receptor 1 precursor - rat
 N:Contains: tumor necrosis factor binding protein 1 (TNF blocking factor)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence_revision 07-Oct-1994 #text_change 02-Jul-1998
 C:Accession: B36555
 R:Hammer, A.; Maurer-Fogy, I.; Kroenke, M.; Scheurich, P.; Pfizenmaier, K.; Lantze, M
 DNA Cell Biol. 9, 705-715, 1990
 A:Title: Molecular cloning and expression of human and rat tumor necrosis factor rece
 A:Reference number: A36555; MUID:91090841
 A:Accession: B36555
 A:Molecule type: mRNA
 A:Residues: 1-461 <HTM>
 A:Cross-references: GB:M63122; NID:g207361; PID:g207362
 C:Comment: This protein is one of two known receptors for both TNF-alpha (cachectin)
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-211/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-201/Product: tumor necrosis factor binding protein #status predicted <TBP>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:212-234/Domain: transmembrane #status predicted <MEM>
 F:235-461/Domain: intracellular #status predicted <INT>
 F:54,151,201/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.1%; Score 129; DB 1; Length 461;
 Best Local Similarity 22.1%; Pred. No. 0.00084;
 Matches 55; Conservative 31; Mismatches 109; Indels 54; Gaps 10;

Oy 6 LPLHRTVFAALLFLHLACKVSCETG-----DCROEFKDRSGNCVLCCKOC 52
 Db 3 LPIVGLLISVLLALLMGHPSGTGLVPSLGDREKRDNLCPGKVAHPKNNISICTKC 62
 Oy 53 GPGMELSKREGCGFGEDACVPCPRHREKDWG-FQCKPCADCAL-VNFORANCSTHS 110
 Db 63 HKGYLVSDC-SPQGEVCEVCCKRGFTTASQNHVROCLCKTRKEMFVEISPCXADM 121
 Oy 111 DAVGDCDLPFGYKRTKLGVFQDMECVPC-----GDPEPP 144
 Db 122 DTVCGCKKNQORILSETHFQCVDCSPFCNCTVTPPCKEKONTYCNCHAGFTLSENETP 181
 Oy 145 YEPHCTSKVNLVKS---STVSSPRDTAAVAVICALATVLLAC---SSCVSSTARSS 197
 Db 182 CS-HCKKNQECMKLCLPVAVNTNPDSGTAVLL-PLVIFLGCLLFFIISILCRPQ 238
 Oy 198 WRNPVASS 206
 Db 239 WR--PRYS 245

RESULT 10

A60771
 B-cell activation protein CD40 precursor - human
 N:Alternate names: B-cell surface antigen Bp50
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 10-Sep-1997

[illegible]

F;22-1786/Product: laminin beta-1 chain #status predicted <MAT>

F:22-270/Domain: VI <DOM6>
 F:271-548/Domain: V <DOM5>
 F:335-395/Domain: laminin-type EGF-like homology <EG1>
 F:398-455/Domain: laminin-type EGF-like homology <EG2>
 F:458-507/Domain: laminin-type EGF-like homology <EG3>
 F:463-468/Region: cell attachment
 F:510-540/Domain: laminin-type EGF-like homology #status atypical <EG5>
 F:549-774/Domain: IV <DOM4>
 F:662-668/Region: cell attachment
 F:773-818/Domain: laminin-type EGF-like homology <EG6>
 F:775-1178/Domain: III <DOM3>
 F:821-864/Domain: laminin-type EGF-like homology <EG7>
 F:867-914/Domain: laminin-type EGF-like homology <EG8>
 F:917-973/Domain: laminin-type EGF-like homology <EG9>
 F:923-927/Region: cell attachment
 F:950-954/Region: cell attachment
 F:976-1025/Domain: laminin-type EGF-like homology <EG10>
 F:1028-1081/Domain: laminin-type EGF-like homology <EG11>
 F:1084-1129/Domain: laminin-type EGF-like homology <EG12>
 F:1132-1136/Domain: laminin-type EGF-like homology <EG13>
 F:1179-1397/Domain: II, heptad repeats <DOM2>
 F:1398-1430/Domain: alpha <ALP>
 F:1431-1786/Domain: I, heptad repeats <DOM1>
 F:30-35/Disulfide bonds: #status predicted
 F:120-356,519,677,965,1041,1195,1279,1336,1343,1487,1542,1643/Binding site: carbohydrate
 F:1179,1182,1785/Disulfide bonds: interchain #status predicted

Query Match 10.8%; Score 125.5; DB 1; Length 1786;
 Best Local Similarity 25.8%; Pred. No. 0.0049;
 Matches 47; Conservative 13; Mismatches 61; Indels 61; Gaps 12;
 QY 10 RTVFAILFLH---LACKVSCETGDCRQEFKDRSGNCV-----LCKQCGPEMEL 58
 Db 756 RNIIF-SISALLHQTGLACE--CDPGSLSSVCDPNNGGQCQCRPNVVGRTCNRCAPGT-- 811
 QY 59 SKEGFGYGEDACQPCRPNR-----FKEDMGFOKC 89
 Db 811 -----FFGFGSG-CKPECHLGQSVNFCNPNVGTGQCHFCQGVARQCDRLCPGHWGFPSC 864
 QY 90 KPCA-----DCALVNRFORANC-SHTSDAYGDCLPFGFYRKLTVGFQDMCEVPCGDDP 142
 Db 865 QPCQGNHADDCCDPVGTG-ECLINCODYTMGNHCERCLAGY-GDPIITGSGD-HCRPCPCPD 921
 QY 143 PP 144
 Db 922 GP 923

RESULT 13
 A26431
 Nerve growth factor receptor precursor, low affinity - rat
 N:Alternate names: NGF receptor
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 19-Nov-1988 #sequence (Norway rat)
 C:Accession: A26431; PH1229
 R:Nadeke, M.J.; Misko, T.P.; Hsu, C.; Herzenberg, L.A.; Shooter, E.M.
 Nature 325, 593-597, 1997
 A:Title: Gene transfer and molecular cloning of the rat nerve growth factor receptor.
 A:Reference number: A26431; M0ID:87115859
 A:Molecule type: mRNA
 A:Residues: 1-425 <RAD>
 A:Cross-references: GB:X05137; NID:956755; PID:956756
 R:Netts, M.; Timmusk, T.; Allikmets, R.; Saarna, M.; Persson, H.
 Gene 121, 247-254, 1992
 A:Title: Regulatory elements and transcriptional regulation by testosterone and retinoid
 A:Reference number: PH1229
 A:Accession: PH1229
 A:Molecule type: DNA
 A:Residues: 1-20 <MET>
 A:Cross-references: GB:X61269

C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastom
 C:Comment: The cysteine-rich region of the extracellular domain may form part or all
 C:Comment: This protein is thought to form a high-affinity receptor when it associate
 C:Genetics: 20/3
 A:Insertions: 20/3
 C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
 C:Keywords: duplication; signal sequence #status predicted <SIG>
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-425/Product: nerve growth factor receptor #status predicted <MAT>
 F:30-251/Domain: extracellular #status predicted <EXT>
 F:33-66/Domain: NGF receptor repeat homology <NG1>
 F:68-109/Domain: NGF receptor repeat homology <NG2>
 F:110-148/Domain: NGF receptor repeat homology <NG3>
 F:150-190/Domain: NGF receptor repeat homology <NG4>
 F:198-249/Region: serine/threonine-rich
 F:252-273/Domain: transmembrane #status predicted <MEM>
 F:274-425/Domain: intracellular #status predicted <INT>
 F:61/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 120.5; DB 2; Length 425;
 Best Local Similarity 20.3%; Pred. No. 0.004;
 Matches 59; Conservative 27; Mismatches 91; Indels 113; Gaps 12;
 QY 17 ILFLHL-----ACKVSCETGDCRQEFKDRSGNCVLCCKQCGPEMELSKGFGYGEDAQ 71
 Db 16 LLLLLILGVSSGAKETCSIG-----LYHSQEC-CKKCNAGEVYAGQCG--ANQIV 64
 QY 72 CVPQPRHFKED--WGFQKPCADCALVNRFORANC-SHTSDAYGDCLPFGFYRKT-- 127
 Db 65 CEPCLDNVTFSDVVSATPEPKPCTEC-LGQSMSPAPCEVDAVVC-RCAYGYQDEETGH 122
 QY 127 -----LVGQDME-----CVPK----- 139
 Db 123 CEACSGVEGSGLVFSCQDKQNTVCEBCEPBGITYSDEANHVDPCLPCVCEDTERQLECT 182
 QY 139 -----GDPPEPPEHCSTKYNLWKISSTVSSPDDTAAV- 173
 Db 183 PMADACEELPGKWRIPSTPEBSSDTAPSTOPEVPPEDDLV--PSTVADMTYTWAGSS 240
 QY 173 -----VICALATVLLACSSGCVSSSTARSGWRNPAVSSH 208
 Db 241 QPVYRGTGNTLIPVYCSTILAAVVGVLVIAFRMNSCKQKNGANSRP 250

RESULT 14
 M0MSB1
 laminin beta-1 chain precursor - mouse
 N:Alternate names: laminin chain B1
 C:Species: Mus musculus (house mouse)
 C>Date: 28-Feb-1986 #sequence (revision 30-Jun-1991 #text change 21-Aug-1998
 C:Accession: A26413; S02679; S05326; S14877; A02871; S02036; S13543
 R:Sasaki, M.; Kato, S.; Kohno, K.; Martin, G.R.; Yamada, Y.
 Proc. Natl. Acad. Sci. U.S.A. 84, 935-939, 1987
 A:Title: Sequence of the cDNA encoding the laminin B1 chain reveals a multidomain pro
 A:Reference number: A26413; M0ID:87147212
 A:Accession: A26413
 A:Molecule type: mRNA
 A:Residues: 1-1786 <SAS>
 A:Cross-references: EMBL:M1552; NID:9198700
 A:Note: translation in Genbank has additional 48 residues at the amino end
 R:Fujitara, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R.
 Biochem. J. 252, 453-461, 1988
 A:Title: Structure and distribution of N-linked oligosaccharide chains on various dom
 A:Reference number: S02678; M0ID:88336259
 A:Accession: S02679
 A:Molecule type: protein
 A:Residues: 28-42;932-946 <FUT>
 R:Hartl, L.; Oberdaemer, I.; Deutzmann, R.
 Eur. J. Biochem. 173, 629-635, 1988
 A:Title: The N terminus of laminin A chain is homologous to the B chains.
 A:Reference number: S00624; M0ID:88225080
 A:Accession: S05326

A:Molecule type: protein
A:Residues: 457-466;654-668;932-946 <HAR>
R:Mann, K.; Deutzmann, R.; Timpl, R.
E: Eur. J. Biochem. 178, 71-80, 1988
A:Title: Characterization of proteolytic fragments of the laminin-nidogen complex and the
A:Reference number: S08895; MUID:89078415
A:Accession: S14877
A:Molecule type: protein
A:Residues: 590-620 <MAN>
R:Barlow, D.P.; Green, N.M.; Kurkinen, M.; Hogan, B.L.M.
E: MBO J. 3, 2355-2362, 1984
A:Title: Sequencing of laminin B chain cDNAs reveals C-terminal regions of coiled-coil a
A:Reference number: A02870; MUID:85051302
A:Accession: A02871
A:Molecule type: mRNA
A:Residues: 1232-1530, 'MBMP', 1535-1691, 'C', 1693-1748, 'N', 1750-1786 <BAR>
A:Cross-references: EMBL:X05212; NID:952861; PID:9809042
R:Deutzmann, R.; Huber, J.; Schmetz, K.A.; Oberbauer, I.; Hartl, L.
E: Eur. J. Biochem. 177, 35-45, 1988
A:Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-ter
A:Reference number: S01790; MUID:89030693
A:Accession: S02036
A:Molecule type: protein
A:Residues: 1561-1587 <DEU>
R:Paulsson, M.; Deutzmann, R.; Timpl, R.; Dalzoppo, D.; Odermatt, E.; Engel, J.
E: EMBO J. 4, 309-310, 1985
A:Title: Evidence for coiled-coil alpha-helical regions in the long arm of laminin.
A:Reference number: S13543; MUID:85257455
A:Accession: S13543
A:Molecule type: protein
A:Residues: 1700-1748, 'N', 1750-1759 <PAU>
C:Genetics:
A:Gene: Lamb-1
A:Map position: 12
C:Complex: laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
C:Function:
A:Description: Interact with cells and with other basement membrane proteins to promote
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
F:1-11/Domain: signal sequence #status predicted <SIG>
F:12-1786/Product: laminin beta-1 chain #status predicted <MAT>
F:22-270/Domain: VI <DOM5>
F:271-540/Domain: V <DOM5>
F:271-332/Domain: laminin-type EGF-like homology <EG1>
F:333-335/Domain: laminin-type EGF-like homology <EG2>
F:398-455/Domain: laminin-type EGF-like homology <EG3>
F:458-507/Domain: laminin-type EGF-like homology <EG4>
F:510-540/Domain: laminin-type EGF-like homology #status atypical <EG5>
F:541-772/Domain: IV <DOM4>
F:773-1182/Domain: III <DOM3>
F:773-818/Domain: laminin-type EGF-like homology <EG6>
F:821-864/Domain: laminin-type EGF-like homology <EG7>
F:867-914/Domain: laminin-type EGF-like homology <EG8>
F:917-973/Domain: laminin-type EGF-like homology <EG9>
F:976-1025/Domain: laminin-type EGF-like homology <EG10>
F:1084-1129/Domain: laminin-type EGF-like homology <EG11>
F:1132-1176/Domain: laminin-type EGF-like homology <EG12>
F:1183-1397/Domain: II, heptad repeats <DOM2>
F:1398-1430/Domain: alpha <ALP>
F:1431-1786/Domain: I, heptad repeats <DOM1>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:302/Disulfide bonds: #status predicted
F:120-356, 519, 677, 1041, 1195, 1279, 1336, 1343, 1487, 1533, 1542, 1643/Binding site: carboxydrat
F:1179, 1182, 1785/Disulfide bonds: interchain #status predicted

Db 756 RNIIH-SISALIHQTLGLACE--CDPQGSLSVCDPNNGGQCQRNRYGRICNRCAPEGT-- 811
OY 59 SKEGFGYGEDACVPCRPBR-----FKEDMGFOKC 89
 ||:| | | |
Db 811 -----FGGGPNG-KKPCRCCHLOGSASAFCAIDAITGCQHCFOGIYARCCDRCLPGWGFPSG 864
OY 90 KPC-----ADCALYNRQRANCSHTSDAVCGDCLDPFYRTKLYGVGDMEVCYSGPPP 143
 ||:| | | | :|:| | | |:
Db 865 QPCQCNHALDCIDVTYGECLSCODPYTGHNCERCLAgy-GDPIIGSGD-HCRPCCPDG 922
OY 144 P 144
 |
Db 923 P 923

RESULT 15

GONUN

nerv growth factor receptor precursor, low affinity - human

N:Alternate names: NGF receptor
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 01-May-1998
C:Accession: A25218; A60204; S21689; I57638
R:Johnson, D.; Lamanah, A.; Buck, C.R.; Sehgal, A.; Morgan, C.; Mercer, E.; Bothwell,
Cell 47, 545-554, 1986
A>Title: Expression and structure of the human NGF receptor.
A:Reference number: A25218; MUID:87051725
A:Accession: A25218
A:Molecule type: mRNA
A:Residues: 1-427 <JOH>
A:Cross-references: GB:M14764; NID:g189204; PID:g189205
J:Marano, N.; Dietzschold, B.; Earley Jr., J.J.; Schattelman, G.; Thompson, S.; Grob,
J. Neurochem. 48, 225-233, 1987
A>Title: Purification and amino terminal sequencing of human melanoma nerve growth fa
A:Reference number: A60204; MUID:87085574
A:Accession: A60204
A:Molecule type: Protein
A:Residues: 29-31,'T','33-42,'TT','45-46,'TX','50-51,'XX','54-56 <MAR>
A:Experimental source: melanoma cell line A875
A>Note: This sequence has been corrected by a note added in proof to follow the nucle
R:Visavajjala, P.; Leszyk, J.D.; Ihn-Goeke, J.; Ross, A.H.
Arch. Biochem. Biophys. 294, 244-252, 1992
A>Title: Structural domains of the extracellular domain of human nerve growth factor
A:Reference number: S21689; MUID:92198017
A:Accession: S21689
A>Status: preliminary
A:Molecule type: protein
A:Residues: 183-208 <VIS>
R:Sehgal, A.; Patil, N.; Chao, M.
Mol. Cell. Biol. 8, 3160-3167, 1988
A>Title: A constitutive promoter directs expression of the nerve growth factor recept
A:Reference number: I57638; MUID:89096903
A:Accession: I57638
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <RES>
A:Cross-references: GB:M21621; NID:g189206; PID:g189207
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastom
C:Comment: The cysteine-rich region of the extracellular domain may form part or all
C:Comment: This protein is thought to form a high-affinity receptor when it associate
C:Comment: This receptor undergoes both N- and O-linked glycosylation.
C:Genetics:
A:Gene: NGFR
A:Cross-references: GDB:120234; OMIM:162010
A:Map position: 17q21-17q22
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor
F:1-78/Domain: signal sequence #status predicted <SIG>
F:29-42/Product: nerve growth factor receptor #status experimental <MAT>
F:29-42/Domain: extracellular #status predicted <XT>
F:32-65/Domain: intracellular repeat homology <NG2>
F:67-108/Domain: NGF receptor repeat homology <NG3>
F:109-147/Domain: NGF receptor repeat homology <NG3>
F:149-189/Domain: NGF receptor repeat homology <NG3>

```
F.197-248/Region: serine/threonine-rich
F.251-272/Domain: transmembrane #status predicted <TRM>
F.273-427/Domain: intracellular #status predicted <INT>
F.60/Binding site: carboxylate (Asn) (covalent) #status predicted
```

Query Match	Similarity	10.2%	Pred. 118.5	DB 1	Length 427
Match	Local Similarity	20.0%	Pred. NO. 0.0059		
Matches	58	Conservative	26	Mismatches	93
				Indels	113
				Gaps	12

QY	17	ILFLHL-----	ACRVSCETGDCROOEFKDRSGNCVLCQKCPGMLSKEGCFGYGEDAQ	71
		:	:	
Db	15	LLLLLLLGVSLGNAEACTG----	LYTHGEC--CRACMLGEBVAPQCG---ANQTV	63
		:	:	
QY	72	CVPCKPHAREKD--	WGFOKCKPCADOCALVNRFORANCSHTSDAVCGDCLPGFYRKTK---	127
		:	:	
Db	64	CEPCLDVSTFSDVVSATBCEKPCPCTEVLGQMS-	APCVADAVNC-RCAYGYQDDETTR	121
		:	:	
QY	127	-----LVGFQDME-----	-----CVPC-----	139
		:	:	
Db	122	CEACVRCVAGSGLVESCQDKQNTVCECPDGYTSDCANHVDPCLPCTVCEQDTERQLRECT		181
		:	:	
QY	139	-----	GDPPEPYEPHCTSKVNLVKLISSITYSSRDRVAAL-	173
		:	:	
Db	182	KWADAECBEIPGRWITRSTPPEGSDSSTASTOEPEAPRODI--	IASTVAGVYTTVMGSS	239
		:	:	
QY	173	-----	VICSAATVLLACSGCSGVSTAGSGWRNRPVASSHP	208
		:	:	
Db	240	QPVYTRGTTDNLIPYCSILAAVVGVAIVIAFKRNNSCQKQNGANSRP		289
		:	:	

Search completed: May 13, 1999, 20:59:04
Job time: 2318 sec

ID 0X40_MOUSE STANDARD; PRT; 272 AA.
 AC P47741;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE 0X40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
 GN TXGP1 OR OX40.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C;
 RX MEDLINE; 94044750.
 RA CALDERHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J.,
 RA CLAASSEN E., NOELLE R.J., FELL H.;
 RL J. IMMUNOL. 151:5261-5271(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 95255413.
 RA BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
 RA BARCLAY A.N.;
 RL EUR. J. IMMUNOL. 25:926-930(1995).
 CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GB34 CYTOKINE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; Z21674; G312828; -
 DR EMBL; X85214; G732819; -
 DR MGD; MGI:104512; TXGP1.
 DR PROSITE; PS00652; TNFR_NGFR_1; 3.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
 KW SIGNAL.
 FT CHAIN 1 19 POTENTIAL.
 FT DOMAIN 20 272 OX40L RECEPTOR.
 FT TRANSEM 20 211 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 212 236 POTENTIAL.
 FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 26 165 4 X TNFR-CYS.
 FT REPEAT 26 61 TNFR-CYS 1.
 FT REPEAT 62 103 TNFR-CYS 2.
 FT REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
 FT REPEAT 125 165 TNFR-CYS 4.
 FT CARBOHYD 144 144 POTENTIAL.
 FT CONFLICT 15 15 A -> G (IN REF. 2).
 SQ SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

Query Match 13.38; Score 155; DB 1; Length 272;
 Best Local Similarity 27.68; Pred. No. 1.4e-06;

Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

QY 18 LFLHACKVSCFEDCKQDFKDRSGN-CYLCKOCGEGMELSECGEGYEDACQVPCR 76
 DB 11 LLLALTLGLVAPRLNCKVHTYP--SGHKC--CRECGGHGMSVRC--DHTRDILCHPCE 64
 QY 77 PHREKEDMGFOCKPCADCAVNRFO--RANCSHTSDAVCGSLGFGYKKTIVGFOD--- 133
 DB 65 TGFNEAVNVTCTCKOCCTOCNHRSGSELKNOCTPQDVTVC--RCRGTOPR-----QDSGY 117
 QY 133 ---MECVPCGDPPEYEP---HCTSKVNLKISSTVSSPDDTVAVAIYC---SALATVL 182
 DB 118 KLGVDCVPC--PQGHFSGGNNQACKPMTNCLSGKQTHRPASDSDLVN--CEDRSILATLL 174
 QY 183 LACS-----SCVSTARGSSWRNPAVSSHPV 210
 DB 175 WETQPTFRPTVOST---TVMPTSTSLSPPTL 205

RESULT 3
 TNRC_HUMAN
 ID TNRC_HUMAN STANDARD; PRT; 435 AA.
 AC P36941;

DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
 GN LTR OR TNFR.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 93252381.
 RA BAENS M., CHAFFANER M., CASSIMAN J.J., DEN BERGHE H., MARYNEN P.,
 RL GENOMICS 16:214-218(1993).
 RN [2]
 RP FUNCTION.
 RX MEDLINE; 94225209.
 RA CROME P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSON C.,
 RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
 RL SCIENCE 264:707-710(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC -1- IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; I04270; G339762; -
 DR HSP; P19999; ICLG.
 DR MIM; 600979; -
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 3.
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 KW SIGNAL.
 FT CHAIN 1 30
 FT DOMAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
 FT TRANSEM 31 227 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 228 248 POTENTIAL.
 FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 42 211 4 X TNFR-CYS.
 FT REPEAT 82 124 TNFR-CYS 1.
 FT REPEAT 125 168 TNFR-CYS 2.
 FT REPEAT 169 211 TNFR-CYS 3.
 FT REPEAT 125 168 TNFR-CYS 4.
 FT DISULFID 43 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 62 80 BY SIMILARITY.
 FT DISULFID 83 98 BY SIMILARITY.
 FT DISULFID 101 116 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT DISULFID 126 132 BY SIMILARITY.
 FT DISULFID 139 148 BY SIMILARITY.
 FT DISULFID 142 167 BY SIMILARITY.
 FT DISULFID 170 185 BY SIMILARITY.
 FT CARBOHYD 40 40 POTENTIAL.
 FT CARBOHYD 177 177 POTENTIAL.
 SQ SEQUENCE 435 AA; 46709 MW; 203B82DD CRC32;

Query Match 12.18; Score 140.5; DB 1; Length 435;
 Best Local Similarity 24.18; Pred. No. 3.4e-05;

Matches 57; Conservative 29; Mismatches 74; Indels 77; Gaps 16;

QY 30 ETGCGROE---FKDRSGNCVLCCKGPGMELSECGEGYEDACQVPCRPHREKEDMGF 86
 DB 39 ENQTCRQDEKEYIEYEPQHTIC--CSRCPGTYVSKC--SHRIDVCTCAENSTNEHWNY 94
 QY 87 ---QCKPC-----ADCAVNRFO-----RANCSHTSDAVCGDCLPGF 121
 DB 95 LTIQLCRPDOPVWGLIELIAPCTSKRTQCRQPGMCAALECTHCE--LLSDCPRT 152
 QY 122 YRTRK-LVGFQDMECVPC-----GDPPEYEPH--CTSKVNLKIS-----STVS 163
 DB 153 EAELEKDEVGKGNHCVCKAGKHFOHTSSPSARCOPHTRCENO--GLVBAADGTAQSDTTCK 211
 QY 164 SPRD-----TAVAVICSLATVTLACS--SCVSTARGSSWRNPAVSSHPV 210

Db 212 NPLELPPEMSTMTMLAVLPLAFLLATVFSCTI-----WK-----SHPSL 254

RESULT 4
VT2_MXVVL STANDARD: PRT: 326 AA.

AC P29825:
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS MYXOMA VIRUS (STRAIN LAUSANNE).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXYIRIDAE; CHORODOPOXYIRINAE;
OC LEBORIPPOXYIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91335768.
RA UPTON C., MACEN J.L., SCHREIBER M., MCFADDEN G.:
RL VIROLOGY 184:370-382(1991).
CC -1- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC EMBL: M95181; G332310; -;
DR EMBL: A23729; E199442; -;
DR PIR: A40566; GOVZML.
DR HSSP: P19438; 1TNR.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW RECEPTOR; GLYCOPROTEIN; REPEAT; SIGNAL.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 326 PROTEIN T2.
FT DOMAIN 27 186 4 X TNFR-CYS.
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 66 66 POTENTIAL.
FT CARBOHYD 181 181 POTENTIAL.
FT CARBOHYD 205 205 POTENTIAL.
FT CARBOHYD 238 238 POTENTIAL.
SQ SEQUENCE 326 AA; 35208 MW; 2F059A61 CRC32;

Query Match 12.0%; Score 140; DB 1; Length 326;
Best Local Similarity 25.0%; Pred. No. 3e-05;
Matches 51; Conservative 20; Mismatches 81; Indels 52; Gaps 12;

QY 13 LFAILLFLHLAC-----KVCETGDCRQDFKDRSGNCVLCQCGPMELSKCGFGY 66
Db 1 MFRLLLLAYAVACVYGAGAPYADRGKCRGNDY-EKDGIC--CTSCPGSYASRLC--GP 55

QY 67 GEDACQVCRPHRFEDMG-QCKPCADCALVNFQRANSHCHTDVACGDCGLPFFY--- 123
Db 56 GSDIVCSCKNETFTASTNHPACVSCGRCTGHLSQSQCDKTRDYC-DOSAGNYCIL 114

QY 123 -----RKTKL-VGF-----QDMECVPCGDEPPPY-----EPHCTSKNVL 155
Db 115 KGQEGCRICAPRTKCPAGYGVSGHRTGVDVCTKC-----PRITYSDAVSSTETCTSENY 170

QY 156 VKISSTVSSPRDTAVAAVICSALA 179
Db 171 ISVEFNLYPVNDTS-----CTTTA 189

RESULT 5
TNRL_MOUSE STANDARD: PRT: 454 AA.
ID TNRL_MOUSE
AC P25118:
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN TNFR1 OR TNFR-1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 9187885.
RA LEWIS M., TARTAGLIA L.A., LEE A., BENNETT G.L., RICE G.C.,
RA WONG G.H., CHEN E.Y., GOEDEL D.V.;
RL PROC. NATL. ACAD. SCI. U.S.A. 88:2830-2834(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91246168.
RA GOODWIN R.G., ANDERSON D., JERZY R., DAVIS T., BRANNAN C.I.,
RA COPELAND N.G., JENKINS N.A., SMITH C.A.;
RL MOL. CELL. BIOL. 11:3020-3026(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91285014.
RA BARRETT K., TAYLOR-FISHWICK D.A., COPE A.P., KISSONERGHIS A.M.,
RA GRAY P.W., FELDMANN M., FOXWELL B.M.J.;
RL EUR. J. IMMUNOL. 21:1649-1656(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE: 92039815.
RA ROTHE J.G., BROCKHAUS M., GENTZ R., LESSLAUER W.;
RL IMMUNOGENETICS 34:338-340(1991).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94245292.
RA BEBO B.F., LINTHICUM D.S.;
RL IMMUNOGENETICS 39:450-451(1994).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93156721.
RA ROTHE J., BLUETEMANN H., GENTZ R., LESSLAUER W., STEINMETZ M.;
RL MOL. IMMUNOL. 30:165-175(1993).
CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -1- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC EMBL: M60468; G199826; -;
DR EMBL: M59377; G302097; -;
DR EMBL: X59238; G53579; -;
DR EMBL: X57796; G54849; -;
DR EMBL: L26349; G430733; -;
DR EMBL: M76656; G202102; -;
DR EMBL: M88067; G202102; JOINED.
DR EMBL: M76655; G202102; JOINED.
DR PIR: A38634; GQKST1.
DR PIR: S16677; S16677.
DR PIR: S19021; S19021.
DR HSSP: P19438; 1TNR.
DR MGD: MGI:98781; TNFR1.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 3.
DR PROSITE: PS50017; DEATH_DOMAIN; 1.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL; APOPTOSIS.

```

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 454 TUMOR NECROSIS FACTOR RECEPTOR 1.
FT DOMAIN 22 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 235 POTENTIAL.
FT DOMAIN 236 454 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 196 4 X TNFR-CYS.
FT REPEAT 43 82 TNFR-CYS 1.
FT REPEAT 83 125 TNFR-CYS 2.
FT REPEAT 126 166 TNFR-CYS 3.
FT REPEAT 167 196 TNFR-CYS 4.
FT DOMAIN 339 349 N-SMASE ACTIVATION DOMAIN (NSD).
FT DOMAIN 356 441 DEATH DOMAIN.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 81 BY SIMILARITY.
FT DISULFID 84 99 BY SIMILARITY.
FT DISULFID 102 117 BY SIMILARITY.
FT DISULFID 105 125 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 146 158 BY SIMILARITY.
FT DISULFID 149 166 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 185 195 BY SIMILARITY.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 202 202 POTENTIAL.
FT CONFLICT 394 394 R -> G (IN REF. 6).
SQ SEQUENCE 454 AA; 50129 MM; 4B6BEC09 CRC32;

```

Query Match 11.9%; Score 138.5; DB 1; Length 454;
Best Local Similarity 22.0%; Pred. No. 5.3e-05;

Matches 56; Conservative 31; Mismatches 103; Indels 65; Gaps 11;

```

QY 6 LPLHRTVLAFLFLHLACKVSCETG-----DCRQOEKDRSGNCVLCQKOC 52
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 3 LPTVPGLLSLVLLALMLMGHPGVTGLVPSLDGREKRDCLCPQKQVHKNNSICCTKC 62

QY 53 GPGMELSKEGFGYGDACVPCPRHREKEDWGF-QKCKFCADCAL-VNFRQANCSHTS 110
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 63 HKGTLYVSDCP-SPGRDTCRECEKGTFTASQVYLRQCLSKTCRKMSQVEISPCQADK 121

QY 111 DAVCG-----DCLPGEFRKTKL-----VGF-QDMECVP 137
   |||||::||::||::||::||::||::||::||::||::||::||::||::||
Db 122 DTYGCGCKENOFQRLSTHQQVCDCSCFENGVTYTPCKETQNTVCNCHAGFFLRESECV 181

QY 138 CGDPPPEPHCTSKVLVVIS-----STVSSPRDTVAVAVICSALATVLLACSSC-VSS 191
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 182 CS-----HCKKNBECMKLCLEPPLANTYNTNQDSTAVLLPLVYLLGLCLLSFIFISL 233

QY 192 TANGSSWRNPVAVS 206
   |||||::||::||::||::||::||::||::||::||::||::||::||::||
Db 234 MCRYPRWR--PEVYS 246

```

RESULT 6

VC22_VARV STANDARD; PRT; 349 AA.

```

AC P34015;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DE 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS VARIOLA VIRUS.
OC VIRIDAE: DS-DNA ENVELOPED VIRUSES; POXYIRIDAE: CHOROPPOXYIRINAE;
OC ORTHOPOXYVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE: 9302281.
RA SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;

```

```

RL FEBS LETT. 319:80-83(1993).
CC -I SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
DR EMBL: X69198; G457087; -.
DR EMBL: X67117; G516449; -.
DR PIR: D36858; D36858.
DR PIR: S35987; S35987.
DR PIR: S46888; S46888.
DR HSSP: p19438; 1TNFR.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW REPEAT.
FT DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MM; 50D0B435 CRC32;

```

Query Match 11.9%; Score 138.5; DB 1; Length 349;
Best Local Similarity 23.7%; Pred. No. 4.2e-05;

Matches 50; Conservative 30; Mismatches 88; Indels 43; Gaps 10;

```

QY 10 RTVLAFLFLHL-----ACKVSCETGDCRQOEKDRSGNCVLCQKOCGPMELSKECG 63
   ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 2 KSVLYLYILFSCIIINGRDAAPYTPNKGCKDTEYK-RHNLG-CISCPPTVSRIC- 58

QY 64 FGYGDAQVPCPRHRE-KEDWGFQKCKFCADCALVNFRQANCSHTSDAVGCDLPGFY 122
   |||||::||::||::||::||::||::||::||::||::||::||::||::||
Db 58 -DSKNTQCTPCGSGTFTSRNNHLPACLSNGRCSNGVETRSCTYHNRIC-ECSPGY 115

QY 123 -----RKTKL-VGF-----QDMECVPG-----DPPPEPHCTSK 152
   ||::||::||::||::||::||::||::||::||::||::||::||::||
Db 116 CLKGSSGCKACYSQKCGIGYVSGHTSVGDIYCSGCGFTYSHTVSSADKCEVPVNT 175

QY 153 VMLKISSTVSSPRDTVAVAVICSALATVLL 183
   ||::||::||::||::||::||::||::||::||::||::||::||::||
Db 176 FNVIDEITLYPVNDISCTRTTGTGSESL 206

```

RESULT 7

NGFR_CHICK STANDARD; PRT; 416 AA.

```

ID NGFR_CHICK
AC P18519;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE (GP80-LNGFR) (P75 ICD).
GN NGFR.
OS GALLUS GALLUS (CHICKEN)
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN.
RX MEDLINE: 90166579.
RA LARGE T.H., WESKAMP G., HEIDER J.C., RADEKE M.J., MISKO T.P.,
RA SHOOTER E.M., REICHARDT L.F.;
RL NEURON 2:1123-1134(1989).
[2]
RP SEQUENCE OF 21-416 FROM N.A.
RX MEDLINE: 90152140.
RA HEUER J.G., FATEMIE-NAINE S., WHEELER E.F., BOTHWELL M.;
RL DEV. BIOL. 137:287-304(1990).
CC -I FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
CC NT-3, AND NT-4.
CC -I SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
CC BOND FORMATION.
CC -I SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -I SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR PIR: JN0006; JN0006.
DR PIR: A60504; A60504.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.

```



```

FT DISULFID 288 313 BY SIMILARITY.
FT DISULFID 315 324 BY SIMILARITY.
FT DISULFID 327 347 BY SIMILARITY.
FT DISULFID 350 359 BY SIMILARITY.
FT DISULFID 352 377 BY SIMILARITY.
FT DISULFID 380 389 BY SIMILARITY.
FT DISULFID 392 410 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 445 441 BY SIMILARITY.
FT DISULFID 443 452 BY SIMILARITY.
FT DISULFID 455 470 BY SIMILARITY.
FT DISULFID 473 487 BY SIMILARITY.
FT DISULFID 475 494 BY SIMILARITY.
FT DISULFID 496 505 BY SIMILARITY.
FT DISULFID 508 522 BY SIMILARITY.
FT DISULFID 784 796 BY SIMILARITY.
FT DISULFID 786 803 BY SIMILARITY.
FT DISULFID 805 814 BY SIMILARITY.
FT DISULFID 817 829 BY SIMILARITY.
FT DISULFID 832 844 BY SIMILARITY.
FT DISULFID 834 851 BY SIMILARITY.
FT DISULFID 853 862 BY SIMILARITY.
FT DISULFID 865 875 BY SIMILARITY.
FT DISULFID 878 887 BY SIMILARITY.
FT DISULFID 880 894 BY SIMILARITY.
FT DISULFID 897 906 BY SIMILARITY.
FT DISULFID 909 925 BY SIMILARITY.
FT DISULFID 928 944 BY SIMILARITY.
FT DISULFID 930 955 BY SIMILARITY.
FT DISULFID 957 966 BY SIMILARITY.
FT DISULFID 969 984 BY SIMILARITY.
FT DISULFID 987 1001 BY SIMILARITY.
FT DISULFID 989 1008 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1023 1036 BY SIMILARITY.
FT DISULFID 1096 1108 BY SIMILARITY.
FT DISULFID 1098 1115 BY SIMILARITY.
FT DISULFID 1117 1126 BY SIMILARITY.
FT DISULFID 1129 1141 BY SIMILARITY.
FT DISULFID 1144 1156 BY SIMILARITY.
FT DISULFID 1146 1163 BY SIMILARITY.
FT DISULFID 1165 1174 BY SIMILARITY.
FT DISULFID 1177 1188 BY SIMILARITY.
FT DISULFID 1191 1191 INTERCHAIN (PROBABLE).
FT DISULFID 1194 1194 INTERCHAIN (PROBABLE).
FT DISULFID 1198 1798 INTERCHAIN (PROBABLE).
FT CARBOHYD 251 251 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 1086 1086 POTENTIAL.
FT CARBOHYD 1250 1250 POTENTIAL.
FT CARBOHYD 1309 1309 POTENTIAL.
FT CARBOHYD 1349 1349 POTENTIAL.
FT CARBOHYD 1500 1500 POTENTIAL.
SQ SEQUENCE 1799 AA; 196352 MW; AF310037 CRC32;

```

Query Match 11.2%; Score 130.5; DB 1; Length 1799;
 Best Local Similarity 23.9%; Pred. No. 0.00081;
 Matches 55; Conservative 22; Mismatches 84; Indels 69; Gaps 14;

```

OY 7 PLHRT-----VLAAILFLHLHACKVCETGDCRQOEFKDRSGNCVLCCKGCGKELSKNE 61
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 761 PLSEFCALILISVSLIYNGALPCQ--CDPOSTLSSECSPHHGQC---RCKPQV-VGRR 813
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 62 C-----GFG-----YGEDACVPCPRHPRF-----KEDMGF 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 814 CDVCAATGYGFGPACQACQCSPDGALSALCEGTSGQC-PCRPAGFGLRCDHCQGGQMGF 872
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 87 QKCRPC-----ADCALVNRFORANC-SHRSDAVCGDCLPGFIRKTKL-VGFQDMQCVPCG 139
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 873 PNCRCVCVNGRDECDTHGACLGRDVTGGEHCRCIAGFGDRLPYGGGCRPC-PCP 931
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 140 DPPPPEPCHTSKVNLIKISSTVSSPRDTAVAAVICSALATVL-LACSSC 188

```

```

Db 932 EGPGSQRHFAI-----SCHRDGYSQQLVCHCRACVITGLRCGAC 969
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 9
ID TNRL_RAT STANDARD; PRT; 461 AA.
AC P22934;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN TNFR1 OR TNFR-1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUKARYOTA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91090841.
RA HIMMLER A., MAUER-FOGY I., KROENKE M., SCHEIDTICH P., PRIZENMAIER K.,
RL LANTZ M., OLSSON I., HAUPTMANN R., STRATOWA C., ADOLF G.R.;
RA DNA CELL BIOL. 9:705-715(1990).
CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -1- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC EMBL; M63122; G207962; -.
DR PIR; B36355; B36355.
DR HSSP; P19438; TNFR.
DR PROSITE; PS00652; TNFR_NGFR.1; 3.
DR PROSITE; PS50050; TNFR_NGFR.2; 3.
DR PROSITE; PS50017; DEATH_DOMAIN: 1.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL; APOPTOSIS.
FT SIGNAL 21
FT CHAIN 1
FT DOMAIN 22 455 TUMOR NECROSIS FACTOR RECEPTOR 1.
FT TRANSMEM 212 234 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 235 461 POTENTIAL.
FT DOMAIN 43 196 CYTOPLASMIC (POTENTIAL).
FT REPEAT 43 82 4 X TNFR-CYS.
FT REPEAT 83 125 TNFR-CYS 1.
FT REPEAT 126 166 TNFR-CYS 2.
FT REPEAT 167 196 TNFR-CYS 3.
FT DOMAIN 344 354 TNFR-CYS 4.
FT DOMAIN 363 448 N-SHASE ACTIVATION DOMAIN (NSD).
FT DISULFID 44 58 DEATH DOMAIN.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 81 BY SIMILARITY.
FT DISULFID 84 99 BY SIMILARITY.
FT DISULFID 102 117 BY SIMILARITY.
FT DISULFID 105 125 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 146 158 BY SIMILARITY.
FT DISULFID 149 166 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 185 195 BY SIMILARITY.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 201 201 POTENTIAL.
SQ SEQUENCE 461 AA; 50969 MW; 82F6B08 CRC32;

```


FT TRANSMEM 183 203 POTENTIAL.
 FT DOMAIN 204 250 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 26 141 3 X TNR-CYS.
 FT REPEAT 64 104 TNR-CYS 1.
 FT REPEAT 105 141 TNR-CYS 2.
 FT CARBOHYD 95 95 TNR-CYS 3.
 FT CARBOHYD 162 162 POTENTIAL.
 SQ SEQUENCE 250 AA; 28164 MW; AF6F74ED CRC32.

Query Match 10.88; Score 125.5; DB 1; Length 250;
 Best Local Similarity 26.9%; Pred. No. 0.00041;
 Matches 49; Conservative 21; Mismatches 59; Indels 53; Gaps 12;

QY 34 CROQEFKDRNGNCVLCQCCGPGMELKECGRGYED---AQCVCRP-HRKEKMGFQ-K 88
 DB 27 CPDHNHTGGGLC--CMWCEGTFEVRDC---EODPTAACDCFCIGTSPDYHTRPH 80
 QY 89 CKPCADCALVNRFORANCSTSDAVCGDCLGFFRKTKLVGFQDMCEVPCG----- 140
 DB 81 CESGRHCN--SGFLIRNCIYTANEC--SC-----SKNOCROECTECDDPLPALTR 130
 QY 140 -----DPPPEYEPYCTSK---VNLVKISSTVSSPRDTAAVAATCSALATVLLACSS 187
 DB 131 QPSETPSPQPPPHLPHTGTEKPSWPLHRLQLPNSTVYSQRSS--HRPLCS-----SD 179
 QY 188 CV 189
 DB 180 CT 181

RESULT 12
 LMB1.HUMAN STANDARD; PRT; 1786 AA.
 AC P07942;
 DT 01-AUG-1988 (REL. 08, CREATED)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
 GN LAMB1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUHERRIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90368768.
 RA VOULTEENAH R., CHOW L.T., TRYGGVASON K.;
 RL J. BIOL. CHEM. 265:15611-15616(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 87280097.
 RA PIKKARAINEN T., EDDY R., FUKUSHIMA Y., BYERS M., SHOWS T.,
 RA PIHLAJANTIE T., SARASTE M., TRYGGVASON K.;
 RL J. BIOL. CHEM. 262:10454-10462(1987).
 RN [3]
 RP SEQUENCE OF 1276-1709 FROM N.A.
 RX MEDLINE: 88021029.
 RA JAYE M., MODI W.S., RICCA G.A., MUDD R., CHIU I.M., O'BRIEN S.J.,
 RA DROHAN W.N.;
 RL AM. J. HUM. GENET. 41:605-615(1987).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC -1- SUBUNIT: THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),
 LAMININ-2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR

CC COMPONENT).
 CC -1- WITH OTHER LAMININ DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 DR EMBL: M61951; G186913; JOINED.
 DR EMBL: M61947; G186913; JOINED.
 DR EMBL: M61918; G186913; JOINED.
 DR EMBL: M61921; G186913; JOINED.
 DR EMBL: M61922; G186913; JOINED.
 DR EMBL: M61923; G186913; JOINED.
 DR EMBL: M61924; G186913; JOINED.
 DR EMBL: M61925; G186913; JOINED.
 DR EMBL: M61926; G186913; JOINED.
 DR EMBL: M61927; G186913; JOINED.
 DR EMBL: M61928; G186913; JOINED.
 DR EMBL: M61929; G186913; JOINED.
 DR EMBL: M61930; G186913; JOINED.
 DR EMBL: M61931; G186913; JOINED.
 DR EMBL: M61932; G186913; JOINED.
 DR EMBL: M61933; G186913; JOINED.
 DR EMBL: M61934; G186913; JOINED.
 DR EMBL: M61935; G186913; JOINED.
 DR EMBL: M61936; G186913; JOINED.
 DR EMBL: M61938; G186913; JOINED.
 DR EMBL: M61939; G186913; JOINED.
 DR EMBL: M61940; G186913; JOINED.
 DR EMBL: M61941; G186913; JOINED.
 DR EMBL: M61942; G186913; JOINED.
 DR EMBL: M61943; G186913; JOINED.
 DR EMBL: M61944; G186913; JOINED.
 DR EMBL: M61945; G186913; JOINED.
 DR EMBL: M61946; G186913; JOINED.
 DR EMBL: M61947; G186913; JOINED.
 DR EMBL: M61948; G186913; JOINED.
 DR EMBL: M61949; G186913; JOINED.
 DR EMBL: M61950; G186913; JOINED.
 DR EMBL: M55370; G186876; JOINED.
 DR EMBL: M55378; G186876; JOINED.
 DR EMBL: M55365; G186876; JOINED.
 DR EMBL: M55371; G186876; JOINED.
 DR EMBL: M55372; G186876; JOINED.
 DR EMBL: M55373; G186876; JOINED.
 DR EMBL: M55374; G186876; JOINED.
 DR EMBL: M55375; G186876; JOINED.
 DR EMBL: M55376; G186876; JOINED.
 DR EMBL: M55344; G186876; JOINED.
 DR EMBL: M55345; G186876; JOINED.
 DR EMBL: M55346; G186876; JOINED.
 DR EMBL: M55347; G186876; JOINED.
 DR EMBL: M55348; G186876; JOINED.
 DR EMBL: M55349; G186876; JOINED.
 DR EMBL: M55350; G186876; JOINED.
 DR EMBL: M55351; G186876; JOINED.
 DR EMBL: M55352; G186876; JOINED.
 DR EMBL: M55353; G186876; JOINED.
 DR EMBL: M55355; G186876; JOINED.
 DR EMBL: M55356; G186876; JOINED.
 DR EMBL: M55357; G186876; JOINED.
 DR EMBL: M55358; G186876; JOINED.
 DR EMBL: M55359; G186876; JOINED.
 DR EMBL: M55360; G186876; JOINED.
 DR EMBL: M55361; G186876; JOINED.
 DR EMBL: M55362; G186876; JOINED.
 DR EMBL: M55363; G186876; JOINED.
 DR EMBL: M55364; G186876; JOINED.
 DR EMBL: M55366; G186876; JOINED.
 DR EMBL: M55367; G186876; JOINED.
 DR EMBL: M55368; G186876; JOINED.
 DR EMBL: M55369; G186876; JOINED.

DR EMBL; M61916; G186837; -;
 DR EMBL; M20206; G186915; -;
 DR PIR; S13547; MMHUB1.
 DR MIM; 150240; -;
 DR PROSITE; PS00022; EGF_1; 9.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
 DR GLYCOPROTEIN; BASEMENT MEMBRANE; EXTRACELLULAR MATRIX; COILED COIL;
 KW LAMININ EGF-LIKE DOMAIN; CELL ADHESION; REPEAT; SIGNAL.
 FT SIGNAL 1 21
 FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.
 FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 271 540 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN V).
 FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 2.
 FT DOMAIN 398 457 LAMININ EGF-LIKE 3.
 FT DOMAIN 458 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 541 771 LAMININ DOMAIN IV.
 FT DOMAIN 772 1178 8 X LAMININ EGF-LIKE REPEATS (DOMAIN III).
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 LAMININ II.
 FT DOMAIN 1398 1430 LAMININ ALPHA.
 FT DOMAIN 1431 1786 LAMININ I.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1353 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1442 1781 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.

FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1179 INTERCHAIN (PROBABLE).
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1785 1785 INTERCHAIN (PROBABLE).
 FT CARBOHYD 120 120 POTENTIAL.
 FT CARBOHYD 356 356 POTENTIAL.
 FT CARBOHYD 519 519 POTENTIAL.
 FT CARBOHYD 677 677 POTENTIAL.
 FT CARBOHYD 1041 1041 POTENTIAL.
 FT CARBOHYD 1195 1195 POTENTIAL.
 FT CARBOHYD 1279 1279 POTENTIAL.
 FT CARBOHYD 1336 1336 POTENTIAL.
 FT CARBOHYD 1343 1343 POTENTIAL.
 FT CARBOHYD 1487 1487 POTENTIAL.
 FT CARBOHYD 1542 1542 POTENTIAL.
 FT CONFLICT 1470 1470 L -> V (IN REF. 3).
 FT CONFLICT 1696 1696 E -> G (IN REF. 3).
 SQ SEQUENCE 1786 AA; 198065 MW; 1D0DB223 CRC32;
 Query Match 10.8%; Score 125.5; DB 1; Length 1786;
 Best Local Similarity 25.8%; Pred. No. 0.0022;
 Matches 47; Conservative 13; Mismatches 61; Indels 61; Gaps 12;
 QY 10 RYVLEALILFLH---LACKYSCETGDCRQDEFKRSGNCV-----LCNQCCPGMEL 58
 Db 756 RNIF-SISLHQTGLACE--CDPQGSLSVCDPENGQCCQCPNVGRTCNRCAPGT-- 811
 QY 59 SKECFGEYGEDACVPCPRHR-----FKEGDFQKC 89
 Db 811 -----FGGPGPS-CRPECCHLDGVSNAFCNPYTGCHCFQGYARQCRCRLPGHMGFSC 864
 QY 90 KPCA-----DCALVNRFORANC-SHTSDAVGDCLPFGYRKTGLVGFQDMCEVPCGDP 142
 Db 865 QPCQNGHADDDPVTG-ELCNCQDYTMGNHCRCLAGY-GDPIIGSD-HCRCPCPD 921
 QY 143 PP 144
 Db 922 GP 923
 RESULT 13
 TNR_MOUSE STANDARD; PRT; 415 AA.
 ID TNR_MOUSE
 AC P50284;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
 GN LTRB OR TNFR.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUHERIA; RODENTIA.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CVB; TISSUE=LUNG;
 RX MEDLINE; 96072804.
 RA FORCE W.R., WALTER B.N., HESSION C., TIZARD R., KOZAK C.A.,
 RA BROWNING J.L., WARE C.F.;
 RL J. IMMUNOL. 155:5280-5288(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE; 96163885.
 RX NAKAMURA T., TASHIRO K., NAZAREA M., NAKANO T., SASAYAMA S.,
 RA HONDO T.;
 RL GENOMICS 30:312-319(1995).
 CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN

CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: U29173; G881621; -
 DR EMBL: U38423; G600223; -
 DR EMBL: U30798; G1061327; -
 DR MGD; MGI:104875; LTR.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00650; TNFR_NGFR_2; 3.
 DR RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 KM SIGNAL 1 30
 FT CHAIN 31 415
 FT DOMAIN 31 223
 FT TRANSMEM 224 244
 FT DOMAIN 245 415
 FT DOMAIN 42 213
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 170
 FT REPEAT 171 213
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 62 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 150
 FT DISULFID 142 169
 FT DISULFID 172 187
 FT CARBOHYD 40 40
 FT CARBOHYD 179 179
 SQ SEQUENCE 415 AA; 44956 MW; 3C5DD121 CRC32;
 Query Match 10.8%; Score 125; DB 1; Length 415;
 Best Local Similarity 22.8%; Pred. No. 0.00069;
 Matches 46; Conservative 24; Mismatches 80; Indels 52; Gaps 9;

QY 47 VLCKCGPGMELEKCEGFGVGEADQVPCPRPHFKEDWG-FQCKPCPADALVNROR-A 104
 DB 57 VCCSRCPGEFVAVC--SRSDTVCTCKPHNSYNEMHLSICOLCPDVLGEFEVA 114
 QY 105 MCSHTSDAVCGDCLPGF-----YRKTLYGFO-----DMECVPC- 139
 DB 115 PCSDKRAEC-RQPGKSCYLDNECVHCEERVLVLCQPTAEVDEIMDTVNCVPC 173
 QY 139 -----GDPPEYEPHCTSKNLVKISTVSSPRDPAVAVICS-----ALATVL 182
 DB 174 PGHONTSSPRACQPH-----TRCEIQGLVEAPGTSYSDTICKNPERGAMLLAILL 228
 QY 183 LACSCVSTARRGSSMRNPAY 204
 DB 229 SLVLFLEFTVLACAMWRHPSL 250

RESULT 14
 WSL1_HUMAN
 ID WSL1_HUMAN STANDARD; PRT; 417 AA.
 AC Q93036; Q93036; Q92983; P78515; Q99831; Q99722; P78507;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DE WSL-1 PROTEIN PRECURSOR (APOPTOSIS-MEDIATING RECEPTOR DR3) (APOPTOSIS-
 MEDIATING RECEPTOR TRAMP) (DEATH DOMAIN RECEPTOR 3) (WSL PROTEIN)
 DE (APOPTOSIS INDUCING RECEPTOR AIR) (APO-3) (LYMPHOCYTE ASSOCIATED
 RECEPTOR OF DEATH) (LARD).
 GN WSL1 OR WSL OR DR3 OR DDR3.
 OS HOMO SAPIENS (HUMAN)
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN (1)
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS.

RC TISSUE-LYMPHOID;
 RX MEDLINE; 97088617.
 RA KITSON J., RAVEN T., JIANG Y.-P., GOEDEL D.V., GILES K.M., PUN K.-T.,
 RA GRINHAM C.J., BROWN R., FARROW S.N.;
 RL NATURE 384:372-375(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE; 97081063.
 RA CHINAIYAN A.M., O'ROURKE K., YU G.-L., LYONS R.H., GARG M.,
 RA DUAN D.R., XING L., GENTZ R., NI J., DIXIT V.M.;
 RL SCIENCE 274:990-992(1996).
 RN (3)
 RP SEQUENCE FROM N.A.
 RA DEGLI-ESPOSTI M.A., DIN W.S., COSMAN D., SMITH C.A., GOODWIN R.G.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
 RN (4)
 RP SEQUENCE FROM N.A.
 RC TISSUE-HEART;
 RA MARSTERS S.A., SHERIDAN J.P., DONAHUE C.J., PITTI R.M., GRAY C.L.,
 RA GODDARD A.D., BAUER K.D., ASHKENAZI A.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
 RN (5)
 RP SEQUENCE FROM N.A.
 RA SCREATOR G., XU X.N., OLSEN A., COMPER A., TAN R., MCMICHAEL A.,
 RA BELL J.;
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
 RN (6)
 RP SEQUENCE OF 4-417 FROM N.A.
 RC TISSUE-BRAIN, AND FETAL LUNG;
 RX MEDLINE; 97205335.
 RA BODMER J.-L., BURNS K., SCHNEIDER P., HOFMANN K., STEINER V.,
 RA THOME M., BORNAND T., HAHNE M., SCHROETER M., WILSON A., FRENCH L.E.,
 RL BROWNING J.L., MACDONALD H.R., TSCHOPE J.;
 RN IMMUNITY 6:79-88(1997).
 RN (7)
 RP SEQUENCE OF 7-417 FROM N.A.
 RC TISSUE-BRAIN;
 RA CHAUDHARY P.M., HOOD L.E.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
 CC -1- FUNCTION: INDICES APOPTOSIS AND ACTIVATES NUCLEAR FACTOR KAPPA-B
 (NF-KAPPA-B). DIRECTLY INTERACTS WITH TRADD ADAPTOR MOLECULE. MAY
 PLAY A ROLE IN REGULATING LYMPHOCYTE HOMEOSTASIS.
 CC -1- SUBUNIT: HOMODIMER. INTERACTS STRONGLY VIA THE DEATH DOMAINS WITH
 THE TNFR1 ASSOCIATED MOLECULE TRADD AND THE TNFR1 RECEPTOR TO
 ACTIVATE AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
 NF-KAPPA B SIGNALING.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMOCYTES AND
 LYMPHOCYTES. DETECTED IN LYMPHOCYTE-RICH TISSUES SUCH AS THYMUS,
 COLON, INTESTINE, AND SPLEEN. ALSO FOUND IN THE PROSTATE.
 CC -1- PTM: GLYCOSYLATED (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS WSL-1/LARD-1A (SHOWN HERE),
 WSL-S1/LARD-3, AND WSL-S2 ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: Y09382; E281462; -
 DR EMBL: Y09382; E281460; -
 DR EMBL: Y09382; E281461; -
 DR EMBL: U72763; G1669512; -
 DR EMBL: U83599; G1800297; -
 DR EMBL: U83600; G1800297; -
 DR EMBL: U78029; G1778764; -
 DR EMBL: U74611; G1763293; -
 DR EMBL: U94501; G2071949; -
 DR EMBL: U94504; G2071955; -
 DR EMBL: U75380; G1695925; -
 DR EMBL: U75381; G1695927; -
 DR EMBL: U83587; G1800293; -
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00650; TNFR_NGFR_2; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW RECEPTOR; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; SIGNAL;
 REPEAT.

Thu May 20 13:49:55 1999

us-09-212-270-3.rsp

Page 12

Db 183 PMADACEEIPGRNIPRSTPEGSDSTAPSTOEVEVPEODLV--PSTVADMTTVKSS 240
QY 173 -----VICSLATVLLACSSCVSTARGSSWRNRNPAVSSH 208
Db 241 QPVYTRGTNDLIPYCSILAAVVGIVAYTAFKRNNSCKOKOGANSRP 290

Search completed: May 19, 1999, 14:43:38
Job time: 45 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:57 ; Search time 48.52 Seconds

(without alignments)
238.778 Million cell updates/sec

Title: US-09-212-270-3

Perfect score: 1162

Sequence: 1 MALKVLPDLHRTVLFALFL.....STANGSSRRNPVSSHPV 210

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPTREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_plant:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	160.5	13.8	348	11	057277	057277 monkeypox v
2	160.5	13.8	348	11	057108	057108 monkeypox v
3	159	13.7	349	11	057291	057291 monkeypox v
4	159	13.7	349	11	057099	057099 monkeypox v
5	159	13.7	349	11	057100	057100 monkeypox v
6	159	13.7	349	11	057101	057101 monkeypox v
7	159	13.7	349	11	057102	057102 monkeypox v
8	156.5	13.5	348	11	057103	057103 monkeypox v
9	152	13.1	267	4	002764	002764 oryctolagus
10	148	12.7	350	11	057116	057116 cowpox viru
11	143.5	12.3	351	11	073559	073559 cowpox viru
12	142.5	12.3	283	2	Q92956	Q92956 homo sapien
13	142	12.2	625	10	035305	035305 mus musculu
14	141.5	12.2	349	11	057110	057110 variola vir
15	141.5	12.2	349	11	057111	057111 variola vir
16	141.5	12.2	349	11	089118	089118 variola vir
17	141.5	12.2	349	11	089098	089098 variola vir
18	141	12.1	355	11	085308	085308 cowpox viru
19	139	12.0	426	2	014865	014865 homo sapien
20	139	12.0	426	2	014865	014865 homo sapien
21	139	12.0	348	11	057112	057112 variola vir
22	139	12.0	348	11	085407	085407 variola vir
23	138.5	11.9	351	11	057117	057117 cowpox viru
24	137.5	11.8	351	11	057121	057121 cowpox viru
25	136.5	11.7	349	11	057284	057284 camelpox vi
26	136.5	11.7	349	11	057098	057098 camelpox vi
27	136	11.7	186	11	072735	072735 cowpox viru
28	135	11.6	349	11	057305	057305 cowpox viru
29	132.5	11.4	380	2	000280	000280 homo sapien

30	132.5	11.4	350	11	057123	057123 cowpox viru
31	130.5	11.2	349	11	057097	057097 camelpox vi
32	130.5	11.2	349	11	057109	057109 variola vir
33	127.5	11.0	326	11	057120	057120 cowpox viru
34	127.5	11.0	326	11	057122	057122 cowpox viru
35	126.5	10.9	347	11	057115	057115 cowpox viru
36	126.5	10.9	360	11	057118	057118 cowpox viru
37	126.5	10.9	347	11	057119	057119 cowpox viru
38	122	10.5	372	2	000279	000279 homo sapien
39	121.5	10.5	418	2	000275	000275 homo sapien
40	121.5	10.5	401	10	070202	070202 mus musculu
41	121	10.4	253	2	000276	000276 homo sapien
42	118.5	10.2	425	2	016042	016042 homo sapien
43	116.5	10.0	401	10	008712	008712 mus musculu
44	115.5	9.9	334	3	Q24403	Q24403 drosophila
45	114.5	9.9	869	12	Q42126	Q42126 xenopus lae

ALIGNMENTS

RESULT 1	
057277	PRELIMINARY; PRT; 348 AA.
AC	01-JUN-1998 (TREMBLREL. 06, CREATED)
DT	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DR	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE	TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN	CRMB.
OS	MONKEYPOX VIRUS.
OC	VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OR	ORTHOPOXVIRUS.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-ZAIRE-1996(96-17), ZAIRES-1996 (96-16);
RA	LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL	SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR	EMBL; U88543; G2738105; -
DR	EMBL; U87841; G2738061; -
DR	PROSITE; PS00652; TNFR_NGFR.1; 2.
SQ	SEQUENCE 348 AA; 38212 MW; 1AABB69 CRC32;
Query Match	13.8%; Score 160.5; DB 11; Length 348;
Best Local Similarity	23.9%; Pred. No. 2e-05;
Matches	56; Conservative 28; Mismatches 71; Indels 79; Gaps 12;
QY	10 RTVLPALFLHLHACKVSC-----ETGDCRQEFKDRSGNCVLCQCGPGM 56
DB	2 RSVLYSYLFL-----SCIIINGRDLPAPHSNGKCKDNEYRSRN--LCCLSCPPT 51
QY	57 ELKSKCGGYGDAQVCRPHRF-KEDMGPKCKPCADCAVYNFQANCSHTSDAVNG 115
DB	52 YASRLC--DSKNTTCTCGSDTFSSHNNHLQACLSGRCSDNOVERSCITNHRIC- 109
QY	116 DCLPFGY-----RKTXL-VGF-----ODMECVPGGDDPPPY----- 146
DB	109 EESPRTYCLLKSSGCRICISITKCGIGVSGYSTGDIVCSPCG--PGTSHVYSSTD 166
QY	146 --EPHCTSKVNLVKTISSTVSSPRDTAAVAVICSALATVLLACSSCVSSTARGSS 197
DB	167 KCEPYSMTFNFDIVEINLYPVNDF-----SCTRTTTLIS 202
RESULT 2	
057108	PRELIMINARY; PRT; 348 AA.
AC	057108;
DT	01-JUN-1998 (TREMBLREL. 06, CREATED)
DR	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DE	01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE	TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.

QY 146 --EPHCTSKVNLVKISSTVSSPRDTAAVAVICSALATVLLACSSCVSSTARSS 197
|| : : : : : ||
Db 167 KCEVSTNFNTFVNDVETINLYPVNDT-----SCRTTTTGLS 202

RESULT 9
ID 002764 PRELIMINARY; PRT; 267 AA.
AC 002764;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE 0440 PRECURSOR (FRAGMENT).
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EIKAROTIA; METAOKA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHB:HM;
RA ISONO T., SETO A.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AB003911; D1020869; -
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PRAM; PF00020; TNFR_c6; 3.
KM SIGNAL.
FT SIGNAL. 1 1
FT NON_TER
FT CHAIN 19 18 POTENTIAL.
FT CHAIN 19 267 OX40.
SQ SEQUENCE 267 AA; 28489 MW; FB3C0C5E CRC32;

Query Match 13.1%; Score 152; DB 4; Length 267;
Best Local Similarity 29.1%; Pred. No. 7e-05;
Matches 62; Conservative 23; Mismatches 100; Indels 28; Gaps 14;
QY 13 LFAAI-LFLHLACKVSC-TGDCRQEFKDRSGNCVLCQCGGMEISGCGFYGEDA 70
|| : : : : : ||
Db 2 LAALGALLLLGLLLGAEPDPDVGDTY---PGGDRCCLECGQGYGVNSRC--NRSDT 56
|| : : : : : ||
QY 71 QCVCRPHREKEDWGFQKRCACADALVNRQ---RANGSHSDAYCGDLPFGYRKT 126
|| : : : : : ||
Db 57 ICHCEGFGFYEAANYQACKCTQC---NRRSGEPQOECTHTDTCV-RRPG---TOP 109
|| : : : : : ||
QY 127 LVGFQ-DMECVPC--GDPPPEPHCTSKVNLVKISSTVSSPRDTAAVAVIC---SALAT 180
|| : : : : : ||
Db 110 LNGXKHGVDACPCQGHSEGNACRPTNCTLAGKRTLDPA-SISDADACEDRSGLAT 168
|| : : : : : ||
QY 181 VLLACSSCV--SSTARGS-SRRNPVAVSSHPV 210
|| : : : : : ||
Db 169 QPWETPSAPYRPTARTSTAMPRTAOGPSTPL 201
|| : : : : : ||

RESULT 10
ID 057116 PRELIMINARY; PRT; 350 AA.
AC 057116;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRAB.
OS COMPOX VIRUS (CPV).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH OPV 89/2(CAT);
RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U90226; G2738129; -
SQ SEQUENCE 350 AA; 38216 MW; 5FF88E39 CRC32;

Query Match 12.7%; Score 148; DB 11; Length 350;
Best Local Similarity 22.9%; Pred. No. 0.00018;
Matches 53; Conservative 27; Mismatches 79; Indels 72; Gaps 11;

QY 10 RTVLFPAIIFLHLACKVSC-----ETGDCRQEFKDRSGNCVLCQCGPGM 56
|| : : : : : ||
Db 2 KSVLYSYILFL-----SCIINGRDVAPYAPSNGCKDNEY-NRHNLC--CLSCPPT 51
|| : : : : : ||
QY 57 ELSEKGFYGEDAOCVPCPHRF-KEDMGFOKRCACADALVNRQRANCSHTSDAVG 115
|| : : : : : ||
Db 52 YARLDSKNTNTQCTPCGSDTFTSRNHLPAQLSCNGRCDSONOVETRSCNTHNRIC- 111
|| : : : : : ||
QY 116 DCLPGFY-----RKTLL-VGF-----QDMECVPCG-----DPPPYE 146
|| : : : : : ||
Db 111 DCAPGYICLLKSGCKACACVSGTCKGIGYVSGHTPTPGDVICSGGIGTYSHTYSSADKE 170
|| : : : : : ||
QY 147 PCHTSKVNLYKISSTVSSPRDTAAVAVICSALATVLLACSSCVSSTARSS 197
|| : : : : : ||
Db 171 PVPSTNFNTFVNDVETINLYPVNDT-----SCRTTTTGLS 203
|| : : : : : ||

RESULT 11
ID 073559 PRELIMINARY; PRT; 351 AA.
AC 073559;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HAR PROTEIN.
GN HAR OR D2L.
OS COMPOX VIRUS (CPV).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GRI-90;
RX MEDLINE; 98229462.
RA SHECHELKUNOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,
RA RAZANKINA O.I., GUTOV V.V., KOTVAL G.J.,
RT "The genomic sequence analysis of the left and right species-specific
RT terminal region of a cowpox virus strain reveals unique sequences and
RT a cluster of intact ORFs for immunomodulatory and host range
RT proteins."
RL VIROLOGY 243:432-460(1998).
DR EMBL; Y15035; E1289276; -
DR EMBL; Y11842; E309299; -
SQ SEQUENCE 351 AA; 28654A80 CRC32;

Query Match 12.3%; Score 143.5; DB 11; Length 351;
Best Local Similarity 22.4%; Pred. No. 0.0004;
Matches 52; Conservative 28; Mismatches 79; Indels 73; Gaps 11;

QY 10 RTVLFPAIIFLHLACKVSC-----ETGDCRQEFKDRSGNCVLCQCGPGM 56
|| : : : : : ||
Db 2 KSVLYSYILFL-----SCIINGRDVAPYAPSNGCKDNEY-NRHNLC--CLSCPPT 51
|| : : : : : ||
QY 57 ELSEKGFYGEDAOCVPCPHRF-KEDMGFOKRCACADALVNRQRANCSHTSDAVG 115
|| : : : : : ||
Db 52 YARLDSKNTNTQCTPCGSDTFTSRNHLPAQLSCNGRCDSONOVETRSCNTHNRIC- 111
|| : : : : : ||
QY 116 DCLPGFY-----RKTLL-VGF-----QDMECVPCG-----DPPPYE 145
|| : : : : : ||
Db 111 ECAPGYICLLKSGCKACACVSGTCKGIGYVSGHTPTPGDVICSGGIGTYSHTYSSADKE 170
|| : : : : : ||
QY 146 EPHCTSKVNLVKISSTVSSPRDTAAVAVICSALATVLLACSSCVSSTARSS 197
|| : : : : : ||
Db 171 EPVPSNTNFNTFVNDVETINLYPVNDT-----SCRTTTTGLS 204
|| : : : : : ||

RESULT 12
ID 092956 PRELIMINARY; PRT; 283 AA.
QY 092956

AC 092956;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HERPESVIRUS ENTRY MEDIATOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: PRIMATES:
CC CATARRHINI: HOMIINAE: HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97053782.
RA MONTGOMERY R.I., WARNER M.S., LUM B.J., SPEAR P.G.;
RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
RL the TNF/NF receptor family";
RL CELL 87:427-436(1996).
DR EMBL: U70321; E319244; -
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PFAM: PF00020; TNFR_C6; 2.
SQ SEQUENCE 283 AA; 30420 MW; B67FBE7B CRC32;

Query Match 12.3%; Score 142.5; DB 2; Length 283;
Best Local Similarity 25.4%; Pred. No. 0.00039;
Matches 60; Conservative 28; Mismatches 83; Indels 65; Gaps 16;
QY 10 RTVFAALLFLHLACKVSCET-GDCRQOEFKDRSGNCVLCQCGPMELSKCGFGYGE 68
DB 17 RTDVLRLVLYLFLAPCAPALPSCKEDEYVGS-EC--CPKCSPPGYRKEACGELTG- 73
QY 69 DAQVPCRPFRKEDW-GFQKCKPCADC--ALVNRFORANCSTHDVAGDCLPFGY--- 123
DB 73 -TVCEPCPPGTIAHLNGSKCLQCMCDPAAGLRASR-NCSTRTENAAGC-CSPGHFCTIV 129
QY 123 -----RKTLYG--FQDMCEVPCGDPPEYEPHCT-----SKVN- 155
DB 130 QDGDHCAACRAVATSPQGRVQKGGESGDTLCQNC--PPGTFSPNGTLEECQHQTCSW 187
QY 155 -LVKISSVSSSR-----DTAVAAVICSALATVLLACSSCVSSTARSSRRNP 202
DB 188 LVTAGAGTSSSHWVWFLSGSLVIVYVCTGLIIT-----CVK-----RRKP 230

RESULT 13
035305 PRELIMINARY; PRT; 625 AA.
AC 035305;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE RECEPTOR ACTIVATOR OF NF-KAPPA (RANK).
GN RANK.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: RODENTIA:
CC SCIURONAHI: MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LIVER EPITHELIUM;
RX MEDLINE: 98032977.
RA ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
RA TOMETSKO M.E., ROUX E.R., TEEPE M.C., DUBOSE R.F., COSMAN D.,
RA GALIBERT L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function";
RL NATURE 390:175-179(1997).
DR EMBL: AF019046; G2612920; -
DR MGD: MGI:1194917; RANK.
DR PFAM: PF00020; TNFR_C6; 3.
SQ SEQUENCE 625 AA; 66621 MW; A01E6C5B CRC32;

Query Match 12.2%; Score 142; DB 10; Length 625;
Best Local Similarity 27.0%; Pred. No. 0.00087;
Matches 54; Conservative 26; Mismatches 76; Indels 44; Gaps 12;

QY 13 LFAALLFLHLACKVSCETGDCRQOEFKDRSGNCVLCQCGPMELSKCGFGYGE 72
DB 15 LIALCVLLVPLQVTLQV--TPPCTQERHXYHLGRG--CSCRCGKYLSSKC--TPTSDDVC 69
QY 73 VPCRPFRKEDWGFQ-KC---KPC-ADCAV-----NFRFORANCSTH-----SDAVC-- 115
DB 70 LFCGPDEYLDTWNEEDKCLHLKRVCDAGAKALVADPGNHTAPRCACCTGAYHNSDCECCR 129
QY 115 --GDCLPGEYRKTLYGFGQDMCEVPC-----GDPPEYEPHCTSKVNLVYKIST 161
DB 130 RNTCEAPFPAGQHPRLQLNKMDVCTPCLLGFFSDVSSSDKCKPM-TNCT-----LLGK 181
QY 162 VSSPRLTAVAVICSALATV 181
DB 182 LEAHGTTESDVYCSSMTL 201

RESULT 14
057110 PRELIMINARY; PRT; 349 AA.
AC 057110;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS VARIOLA VIRUS.
OC VIRUSES: DSDNA VIRUSES, NO RNA STAGE: POXYIRINAE; CHORDOROXVIRINAE;
CC ORTHOROXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MAJOR, CONGO-1970 (CNG-70);
RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U88147; G2738093; -
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
SQ SEQUENCE 349 AA; 38273 MW; BEC101E8 CRC32;

Query Match 12.2%; Score 141.5; DB 11; Length 349;
Best Local Similarity 24.2%; Pred. No. 0.00056;
Matches 51; Conservative 30; Mismatches 87; Indels 43; Gaps 10;
QY 10 RTVFAALLFLHL-----ACKVSCETGDCRQOEFKDRSGNCVLCQCGPMELSKCG 63
DB 2 KSVLYLYLFLSCILINGRDAPYTPPNKCKDTEYK-RHNC--CLSCPCTYASRLC- 58
QY 64 FGYGDAQCVPCRPFRF-KEDWGFQKCKPCADCALVNRFORANCSTHDVAGDCLPFGY 122
DB 58 -DSKTNCTCPCSGSTFTSRNHLPACLSCNCRCSNNGVETRSCVTHTNRIC-ECSPGY 115
QY 123 -----RKTLY-VGF-----QDMCEVPCGDPPEY-----EPHCSTK 152
DB 116 CLKGSSGCKACVSGTQKIGIGVSGHTSVGBVYISPCGFGTYSVSADKCEVPNNT 175
QY 153 VNLVKISSTVSSPRTAAVAVICSALATVYL 183
DB 176 FVIVDEYTLIVVNDTSCRTITTTLSIL 206

RESULT 15
057111 PRELIMINARY; PRT; 349 AA.
AC 057111;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS VARIOLA VIRUS.
OC VIRUSES: DSDNA VIRUSES, NO RNA STAGE: POXYIRINAE; CHORDOROXVIRINAE;
CC ORTHOROXVIRUS.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MINOR, SIERRA LEONE-1968 (SLN-68);
RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U88150; G2738099; -;
DR PROSITE: PS00652; TNER_NGFR_1; 2.
SQ SEQUENCE 349 AA; 38209 MW; 8DAC3FB5 CRC32;

Query Match 12.2%; Score 141.5; DB 11; Length 349;
Best Local Similarity 24.2%; Pred. No. 0.00056;
Matches 51; Conservative 30; Mismatches 87; Indels 43; Gaps 10;

QY 10 RTVLFAIIFLHL-----ACKVSCETGDCRQEFDRSGNCVLCKQGGPMGELSKECG 63
DB 2 KSVLYLXILFLSCIIINGRDAAPYTPPENGCKDTEYK-RHNLG-CLSCPPTIYASRLC- 58
QY 64 FGYGDAQCVPQPRHPR-KEDWGFQCKPCADCALVNRFORANCSTSDAVGDCDLPFGY 122
DB 58 -DSKTNTOCTPCGSGFTSRNNHLPACLSGNGRONSQVETRSNTTHNRIC-ECSFGY 115
QY 123 -----RKTKL-VGF-----QDMECVPCGDPDPY-----EPHCTSK 152
DB 116 CLKSSGCKACYSQYKCGIGYGVSGHTSVGDYICSPCGFTYSYTVSSTDKCEPVNNT 175
QY 153 VNLKISSTVSPRDTAVAAVICALATVLL 183
DB 176 FNYIDVEITLYPYNDISCTRTTGTGLSEITL 206

Search completed: May 14, 1999, 21:39:58
Job time: 7992 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 17:21:18 ; Search time 42.79 seconds

(without alignments)
71.372 Million cell updates/sec

Title: US-09-212-270-4

Perfect score: 873
Sequence: 1 MALKVLEPHRTVLFALILFL.....QDMECVPCGDPPEPPHPC 151

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database: A_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	714.5	81.8	417	1	Amino acid sequenc
2	714.5	81.8	423	1	Amino acid sequenc
3	135.5	13.5	438	1	Plasmod pDC406/OX4
4	135.5	13.5	206	1	Mouse type-II memb
5	135.5	15.5	438	1	OX40/Fc mutein. Pu
6	135.5	15.5	206	1	Mouse OX40 extrac
7	129.5	14.8	833	1	Human DR3 protein.
8	129	14.8	625	1	Murine NF-kB recep
9	129	14.8	616	1	NF-kB receptor act
10	129	14.8	451	1	RANK polypeptide p
11	129	14.8	451	1	RANK polypeptide p
12	129	14.8	616	1	NF-kB receptor act
13	129	14.8	625	1	Murine NF-kB recep
14	128.5	14.7	283	1	Human tumour neuro
15	128.5	14.7	283	1	Herpesvirus entry
16	128	14.7	616	1	A receptor protein
17	127.5	14.6	283	1	Human herpes simpl
18	127	14.5	466	1	Human lamelin B1 c
19	126	14.4	186	1	Cowpox virus psc 1
20	126	14.4	1786	1	Human lamelin B1 c
21	125.5	14.4	283	1	Human TNF receptor
22	125	14.3	197	1	Extracellular doma
23	124.5	14.3	355	1	Cowpox virus T2-eg
24	123.5	14.1	417	1	Death domain conta
25	123.5	14.1	428	1	Death domain conta
26	123.5	14.1	181	1	Human apoptosis pr
27	123.5	14.1	417	1	Human apoptosis pr
28	123.5	14.1	417	1	Human apoptosis pr
29	120	13.7	1764	1	Primary amino acid
30	120	13.7	1776	1	Mouse lamelin B1 c
31	119	13.6	176	1	Mouse lamelin B1 c
32	118	13.5	411	1	Mouse apoptosis in
33	117	13.4	277	1	Deduced sequence e
34	116	13.3	277	1	ACT-4 cell surface
35	116	13.3	277	1	ACT-4-h-1 receptor
36	115	13.2	326	1	Myxoma virus T2 pr
37	115	13.2	326	1	Myxoma virus T2 pr
38	114.5	13.1	277	1	CD40 protein. CD40
39	111.5	12.8	591	1	RANK polypeptide p
40	111.5	12.8	591	1	RANK polypeptide p
41	110.5	12.7	1798	1	Human lamelin B2 c
42	110	12.6	260	1	Human CD27 antigen
43	109.5	12.5	260	1	T lymphocyte-speci

44 108.5 12.4 255 1 R64197 Human 4-1BB polype
45 108.5 12.4 255 1 R70977 H4-1BB receptor pr

ALIGNMENTS

RESULT 1
ID W70386 standard; Protein; 417 AA.
AC W70386;
DE 02-DEC-1998 (first entry)
KW Human; alpha-OAF065; stroma cell; antibody; inflammatory;
OS cytokine-mediated disease; rheumatism; ulcerative colitis.
FH Key Location/Qualifiers
FT Misc-difference 223
FT Misc-difference 224 /note- "encoded by AGA"
FT Misc-difference 224 /note- "encoded by CCR"
FT W09838304-A1.
PD 03-SEP-1998.
PE 26-FEB-1998; J00799.
PR 27-FEB-1997; JP-043143.
PA (ONOX) ONO PHARM CO LTD.
PI Fukushima D, Konishi M, Tada H;
DR WPI: 98-481205/41.
DR N-PSDB: V33361.
PT Membrane polypeptide expressed by human stroma cells, and antibodies
PT recognising it - for treatment of inflammatory and other
PT cytokine-mediated diseases.
PS Claim 1; Pages 28-30; 54pp; Japanese.
CC This is the amino acid sequence of the human alpha-OAF065, used in
CC the method of the invention. The process involves the use of peptides
CC expressed by stroma cells, and its antibodies are used for in the
CC prevention and treatment of inflammatory and other cytokine-mediated
CC diseases such as rheumatism, ulcerative colitis.
SQ Sequence 417 AA;

Query Match 81.8%; Score 714.5; DB 1; Length 417;
Best Local Similarity 82.0%; Pred. No. 2.9e-57;
Matches 123; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

QY 1 MALKVLEPHRTVLFALILFLHACKVSCERTGDSROEFDRSGNCVLCCKGPGMELS 60
DB 1 MALKVLEPHRTVLFALILFLHACKVSCERTGDC-RDOEFDRSGNCVPCNQGMELS 59
QY 61 KECGFGYGEDAQVPCPRHREKEDMGFOCKPCADCALVNRFORANCSHTSDAFCGDCLP 120
DB 60 KECGFGYGEDAQVPCPRHREKEDMGFOCKPCADCAVNVNFORANCSATSDAFCGDCLP 119
QY 121 GFYRKTKLVGFQDMECVPCGDPPEPPHPC 150
DB 120 GFYRKTKLVGFQDMECVPCGDPPEPPHPC 149

RESULT 2
ID W70387 standard; Protein; 423 AA.
AC W70387;
DE 02-DEC-1998 (first entry)
KW Human; beta-OAF065; stroma cell; antibody; inflammatory;
OS cytokine-mediated disease; rheumatism; ulcerative colitis.
FH Key Location/Qualifiers
FT Misc-difference 223
FT Misc-difference 224 /note- "encoded by AGA"
FT Misc-difference 224 /note- "encoded by CCR"
FT W09838304-A1.

PD 03-SEP-1998.
 PF 26-FEB-1998; J00799.
 PR 27-FEB-1997; JP-043143.
 PA (ONO) ONO PHARM CO LTD.
 PI Fukushima D, Konishi M, Tada H;
 DR WPI: 98-481205/41.
 DR N-PSDB: V33362.
 PT Membrane polypeptide expressed by human stroma cells, and antibodies
 PT recognizing it - for treatment of inflammatory and other
 PT cytokine-mediated diseases.
 PS Disclosure: Pages 37-49; 54pp; Japanese.
 CC This is the amino acid sequence of the human beta-OAF065, used in
 CC the method of the invention. The process involves the use of peptides
 CC expressed by stroma cells, and its antibodies are used for in the
 CC prevention and treatment of inflammatory and other cytokine-mediated
 CC diseases such as rheumatism, ulcerative colitis.
 SQ Sequence 423 AA.

Query Match 81.8%; Score 714.5; DB 1; Length 423;
 Best Local Similarity 82.0%; Pred. No. 36-57; Mismatches 17; Indels 1; Gaps 1;
 Matches 123; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

OY 1 MALKVLPHRTVLAAILFLHLACKVSCETGDCSRQSEFKDRSGKCYLCKQCGPEMELS 60
 DB 1 MALKVLLEGEKTEFTLLVLLGLYSKVCETGDC-RQSEFRDRSGKCYPCNCGPEMELS 59
 OY 61 KCGCGFYGDACVPCPRPRFRKEDMGFOCKRACALVNRGRANCSTSDAVGDCDLP 120
 DB 60 KCGCGFYGDACVCTRLRFRKEDMGFOCKRCPCLDAVNRGRANCSTSDAICDCLP 119
 OY 121 GFYRKTLYGFQDMCEVPCGDPPEPPPHC 150
 DB 120 GFYRKTLYGFQDMCEVPCGDPPEPPPHC 149

RESULT 3
 ID R81882 standard; Protein: 438 AA.
 AC R81882;
 DE 30-MAR-1996 (first entry)
 DE Plasmid pDC406/OX40/Fc* encoding an OX40/Fc mutein protein.
 KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;
 KM pDC406/OX40/Fc*; membrane glycoprotein.
 OS Synthetic.
 PN US5457035-A.
 PD 10-OCT-1995.
 PF 23-JUL-1993; 097827.
 PR 23-JUL-1993; US-097827.
 PA (IMMUNEX) IMMUNEX CORP.
 PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
 DR WPI: 95-357992/46.
 DR N-PSDB: T00826.
 PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
 PT and host cells, used to produce recombinant ligand used in e.g.
 PT prim. T cell culture, to modulate immune response etc.
 PS Example 2: Column 35-38; 26pp; English.
 CC This plasmid encodes an OX40/Fc antibody fragment mutein protein.
 CC and is used to express a soluble OX40/Fc mutein fusion protein for
 CC use in detecting CDNA clones encoding a OX40 ligand. The Fc
 CC fragment may be derived from human IgG1, and the plasmid may be
 CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
 CC line. Culture supernatant was purified by affinity chromatography
 CC and this was used, together with labeled goat anti-human IgG to
 CC screen various cell lines.
 SQ Sequence 438 AA.

Query Match 15.5%; Score 135.5; DB 1; Length 438;
 Best Local Similarity 29.4%; Pred. No. 4.3e-05;
 Matches 35; Conservative 14; Mismatches 43; Indels 27; Gaps 7;
 OY 46 NCV-----LCKQCGPMELSKCEGFGYGDACVPCPRPRFRKEDMGFOCKRCACD 96

DB 26 NCVKHTYPSGKHCKCEQCPGHGMVNR--DHTRDYLCHPCETGEFYNENAVNDTCQCTQC 83
 OY 97 ALVNRFO-RANCSHTSDAVGDCDLPGFYRKTLYGFQD-----MECVPCGDPPEPPH 148
 DB 84 NHRSGSELKQNTPTQDYVC-RCRPGTQPR-----QDSGYKLGVDCVPC--PRGHFSP 133

RESULT 4
 ID R81881 standard; Protein: 206 AA.
 AC R81881;
 DE 08-JUL-1996 (first entry)
 DE Mouse type-II membrane polypeptide OX40 extracellular domain.
 KW OX40; OX40-L; cytokine; cell surface molecule;
 KW membrane glycoprotein.
 OS Mus musculus.
 PN US5457035-A.
 PD 10-OCT-1995.
 PF 23-JUL-1993; 097827.
 PR 23-JUL-1993; US-097827.
 PA (IMMUNEX) IMMUNEX CORP.
 PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
 DR WPI: 95-357992/46.
 DR N-PSDB: T00826.
 PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
 PT and host cells, used to produce recombinant ligand used in e.g.
 PT prim. T cell culture, to modulate immune response etc.
 PS Example 1; Column 33-34; 26pp; English.
 CC This sequence encodes the extracellular domain of OX40, a membrane
 CC glycoprotein present on the CD4 positive subset of activated T
 CC cells.
 SQ Sequence 206 AA.

Query Match 15.5%; Score 135.5; DB 1; Length 206;
 Best Local Similarity 29.4%; Pred. No. 2.1e-05;
 Matches 35; Conservative 14; Mismatches 43; Indels 27; Gaps 7;

OY 46 NCV-----LCKQCGPMELSKCEGFGYGDACVPCPRPRFRKEDMGFOCKRCACD 96
 DB 26 NCVKHTYPSGKHCKCEQCPGHGMVNR--DHTRDYLCHPCETGEFYNENAVNDTCQCTQC 83
 OY 97 ALVNRFO-RANCSHTSDAVGDCDLPGFYRKTLYGFQD-----MECVPCGDPPEPPH 148
 DB 84 NHRSGSELKQNTPTQDYVC-RCRPGTQPR-----QDSGYKLGVDCVPC--PRGHFSP 133

RESULT 5
 ID W48976 standard; Protein: 438 AA.
 AC W48976;
 DE 25-SEP-1998 (first entry)
 DE OX40/Fc mutein.
 KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;
 KW chimeric.
 OS Chimeric - Homo sapiens.
 FH Chimeric - Mus sp.
 FH Key Location/Qualifiers
 FT Region 1..206
 FT Region /note="Extracellular domain of mouse OX40"
 FT Region 207..438
 FT Region /note="Mutant Fc region of human IgG1 antibody"
 FT Misc_difference 225
 FT /note="changed from Leu in wild-type to Ala in
 FT mutant"
 FT Misc_difference 226
 FT /note="changed from Leu in wild-type to Gly in
 FT mutant"
 FT Misc_difference 228
 FT /note="changed from Gly in wild-type to Ala in
 FT mutant"
 PN US5783665-A.

KW RANKL; RANK ligand; tumour necrosis factor; TNF; murine.
OS Mus sp.
PN W09828426-A2.
PD 02-JUL-1998.
PE 22-DEC-1997; U23775.
PR 14-OCT-1997; US-064671.
PR 23-DEC-1996; US-059978.
PR 07-MAR-1997; US-813509.
PA (IMMUNEX CORP.
PI Anderson DM, Galibert LJ, Maraskovsky E;
DR WPI: 98-377657/32.
DR N-PSDB: VA1379.
PT New isolated ligand for receptor activator of NF-kappa B - used to
PT develop products for augmenting an immune response for inhibiting an
PT inflammatory response and for protection of cells
PS Example 14; Pages 62-64; 80pp; English.
CC This represents the murine mURANK (receptor activator of necrosis
CC factor-kappaB (NF-kB)) polypeptide which is a homolog of the human RANK.
CC RANK is a member of the tumour necrosis factor (TNF) family. A soluble
CC expressing membrane-associated RANK with a soluble RANK which binds to
CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be
CC used to induce maturation of dendritic cells and enhance their
CC allo-stimulatory capacity, thereby augmenting an immune response. The
CC soluble RANK polypeptide composition may also be used for regulating an
CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists
CC may be useful in ameliorating negative effects of an inflammatory
CC response that result from triggering of RANK, e.g. in treating toxic
CC shock or sepsis, graft-versus-host reactions, or acute inflammatory
CC reactions. They can also be used in adjunct therapy for disease
CC characterised by neoplastic cells that express RANK. RANKL polypeptides
CC can also be used to identify inhibitors of RANK and thus inhibitors of
CC an inflammatory response, and also for protecting RANK-expressing cells
CC from the negative effects of chemotherapy or the presence of high levels
CC of TNF-alpha. The products can also be used for detection and drug
CC screening.
SQ Sequence 625 AA;

Query Match 14.8%; Score 129; DB 1; Length 625;
Best Local Similarity 29.9%; Pred. No. 0.00023;
Matches 44; Conservative 20; Mismatches 57; Indels 26; Gaps 10;

QY 13 LPAALFLHLACKYSCETGDCSRQOEFKDRSGNCVLCQCGPMELSKCGFGYGEDAQ 72
DB 15 LALCVLLPLQVLTQV--TPCTGRHY-EHLGRG--CSRCEPGRYSSKC--TPTSQV 68
QY 73 CVPCHPRFKEDMGFO-KC---KPC-ADCALY-----NRFQANCSHT-----SDAVC- 116
DB 69 CLPCGPDEYLDTWNEEDCKLHKVCDAGKALVADPDGNHTAPRCACAGYHMNSDCBCC 128
QY 116 --GDCLPGFYRKTGVGFQDMCEVCP 139
DB 129 RRMTECARGFGAHPLOJLNKDTVCIPC 155

RESULT 9
W69952
ID W69952 standard; Protein: 616 AA.
AC W69952;
DT 08-OCT-1998 (first entry)
DE NF-kB receptor activator RANK.
KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
KW immune response; inflammatory response; toxic shock; sepsis;
KW RANKL; RANK ligand; tumour necrosis factor; TNF.
OS Homo sapiens.
PN W09828426-A2.
PD 02-JUL-1998.
PE 22-DEC-1997; U23775.
PR 14-OCT-1997; US-064671.
PR 23-DEC-1996; US-059978.
PR 07-MAR-1997; US-813509.
PA (IMMUNEX CORP.
PI Anderson DM, Galibert LJ, Maraskovsky E;
DR WPI: 98-377657/32.
DR N-PSDB: VA1379.
PT New isolated ligand for receptor activator of NF-kappa B - used to
PT develop products for augmenting an immune response for inhibiting an
PT inflammatory response and for protection of cells

PI Anderson DM, Galibert LJ, Maraskovsky E;
DR WPI: 98-377657/32.
DR N-PSDB: VA1379.
PT New isolated ligand for receptor activator of NF-kappa B - used to
PT develop products for augmenting an immune response for inhibiting an
PT inflammatory response and for protection of cells
PS Example 1; Pages 48-51; 80pp; English.
CC This represents the polypeptide RANK (receptor activator of necrosis
CC factor-kappaB (NF-kB)). RANK is a member of the tumour necrosis factor
CC (TNF) family. A soluble RANK may be used for inhibiting activation of
CC NF-kB, by contacting a cell expressing membrane-associated RANK with a
CC soluble RANK which binds to RANK ligand (RANKL). RANKL polypeptides can
CC activate RANK and can be used to induce maturation of dendritic cells and
CC enhance their allo-stimulatory capacity, thereby augmenting an immune
CC response. The soluble RANK polypeptide composition may also be used for
CC regulating an immune or inflammatory response. Inhibition of NF-kB by
CC RANK antagonists may be useful in ameliorating negative effects of an
CC inflammatory response that result from triggering of RANK, e.g. in
CC treating toxic shock or sepsis, graft-versus-host reactions, or acute
CC inflammatory reactions. They can also be used in adjunct therapy for
CC disease characterised by neoplastic cells that express RANK. RANKL
CC polypeptides can also be used to identify inhibitors of RANK and thus
CC inhibitors of an inflammatory response, and also for protecting RANK-
CC expressing cells from the negative effects of chemotherapy or the
CC presence of high levels of TNF-alpha. The products can also be used for
CC detection and drug screening.
SQ Sequence 616 AA;

Query Match 14.8%; Score 129; DB 1; Length 616;
Best Local Similarity 21.5%; Pred. No. 0.00023;
Matches 45; Conservative 28; Mismatches 60; Indels 76; Gaps 12;

QY 7 PLHRVFLAALFLHLACKYSCETGDCSRQOEFKDRSGNCVLCQCGPMELSKCGFG 66
DB 10 PLFALLLCALRLQVALQIA--PCTSEKHV-EHLGRG--CNKCEPGRYSSKC--T 61
QY 67 XGEDACVCPCHPRFKEDMGFO-KC-----KPCA----- 95
DB 62 TTSDSVCLPCGPDYLDWNEEDCKLHKVCDTGKALVAVAGNSTPBRCACTAGYHWS 121
QY 95 -DCALVNRFORANCS-----H---TSDAVCGDCLPGFY----- 124
DB 122 QDCECCRR--NTECARGAHPLOJLNKDTYCKPCLAFYDFDAFSSDTKCRPWTCFLG 179
QY 124 RKTGVGFQDMCEVCPGDP--PPEBPH 149
DB 180 KRVEHGTREKSDAVCSSLPARKPNEPH 208

RESULT 10
W69951
ID W69951 standard; Protein: 451 AA.
AC W69951;
DT 08-OCT-1998 (first entry)
DE RANK polypeptide partial sequence encoded by cDNA clone 9D-15C.
KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
KW immune response; inflammatory response; toxic shock; sepsis;
KW RANKL; RANK ligand; tumour necrosis factor; TNF.
OS Homo sapiens.
PN W09828426-A2.
PD 02-JUL-1998.
PE 22-DEC-1997; U23775.
PR 14-OCT-1997; US-064671.
PR 23-DEC-1996; US-059978.
PR 07-MAR-1997; US-813509.
PA (IMMUNEX CORP.
PI Anderson DM, Galibert LJ, Maraskovsky E;
DR WPI: 98-377657/32.
DR N-PSDB: VA1379.
PT New isolated ligand for receptor activator of NF-kappa B - used to
PT develop products for augmenting an immune response for inhibiting an
PT inflammatory response and for protection of cells

PS Example1: Pages 44-46: 80pp. English.
CC This represents a partial sequence of the polypeptide RANK (receptor
CC activator of necrosis factor- κ apab (NF- κ B)). RANK is a member of
CC the tumour necrosis factor family. A soluble RANK may be used for
CC inhibiting activation of NF- κ B, by contacting a cell expressing membrane
CC -associated RANK with a soluble RANK which binds to RANK ligand (RANKL).
CC RANKL polypeptides can activate RANK and can be used to induce maturation
CC of dendritic cells and enhance their allo-stimulatory capacity, thereby
CC augmenting an immune response. The soluble RANK polypeptide composition
CC may also be used for regulating an immune or inflammatory response.
CC Inhibition of NF- κ B by RANK antagonists may be useful in ameliorating
CC negative effects of an inflammatory response that result from triggering
CC of RANK, e.g. in treating toxic shock or sepsis, graft-versus-host
CC reactions, or acute inflammatory reactions. They can also be used in
CC adjunct therapy for disease characterised by neoplastic cells that
CC express RANK. RANKL polypeptides can also be used to identify inhibitors
CC of RANK and thus inhibitors of an inflammatory response, and also for
CC protecting RANK-expressing cells from the negative effects of
CC chemotherapy or for the presence of high levels of TNF- α . The products
CC can also be used for detection and drug screening.
50 Sequence 451 AA;

Query Match	14.8%	Score	129	DB 1	Length	451
Best Local Similarity	21.5%	Pred. No.	0.00017			
Matches 45	Conservative	28	Mismatches	60	Indels	76
					Gaps	12

QY 7 PLHRTVFAAILFLLHLACKVSCETGDCSRQQEFKDRSGNCVLCCKQCGPMELSKCEGFG 66

Db 10 PLFALLLCALRLQVALQIA--PPCTSEKHY-EHLGRC--CNKCEPGKYMSSKC--T 61

QY 67 YGEDAQCVP CRPHRFKEDWGFQ-KC-----KPCA----- 95

Db 62 TTSDSVCLPCGPDEYLD SWNEEDKCLLHKVCDTGKALVAVVAGNSTTPRCAGTAGYHWS 122

QY 95 -DCALVNRFRANC-----H---TSDAVCGDCLPGFY----- 12

Db 122 QDCECCR--NTECAPGLGAQHPLQLNKDVTCKPCLAGYFSDAFSSTDCKRPWTNCTFLG 179

QY 124 RKT^{KL}VG^{FQ}DM^{EC}VP^{CG}DP⁻⁻⁻PRYEPH 149

Db 180 KVEHHGTEKSDAVCSSLPARKPPNEPH 208

RESULT	11
W68287	
ID	W68287 standard; Protein; 451 AA

PI New isolated receptor activator of necrosis factor-kappa B - useful
PI for, e.g. developing products for regulating an immune or
PI inflammatory response, treating toxic shock or sepsis
PS Example 1: Pages 44-46: 80pp: English
CC This represents a partial sequence of the polypeptide RANK (receptor
CC activator of necrosis factor-kappaB (NF-kB)). RANK is a member of
CC the tumour necrosis factor family. Host cells transformed or transfected
CC with an expression vector comprising the RANK encoding nucleic acid can
CC be used to produce recombinant RANK protein. The soluble RANK may be used

for inhibiting activation of NF- κ B, by contacting a cell expressing membrane-associated RANK with a soluble RANK which binds to RANK ligand (RANKL). The soluble RANK polypeptide composition may also be used for regulating an immune or inflammatory response. Inhibition of NF- κ B by RANK antagonists may be useful in ameliorating negative effects of an inflammatory response that result from triggering of RANK, e.g. in treating toxic shock or sepsis, graft-versus-host reactions, or acute inflammatory reactions. They can also be used in adjunct therapy for disease characterised by neoplastic cells that express RANK. The products can also be used for detection and drug screening.

Query Match	14.8%	Score 129	DB 1	length 451
Best Local Similarity	21.58%	Pred. No. 0.00017		
Matches 45	Conservative 28	Mismatches 60	Indels 76	Gaps 12

QY 7 PLHRTVFAAILFLLHLACKVSCETGDCSRQQEFKDRSGNCVLCKQCGPGMELSKCEGFG 66

Db 10 PLFALLLCALRLQVALQIA--PRCTSEKHY-EHLGRC--CNKCEPGKYMSSKC--T 61

0Y 67 YGEDAQCVPICRPHREKEDWGFQ-KC-----KPCA----- 95

Db 62 TTSDSVCLPCGPDEYLD SWNEEDKCLHKVCDTGKALVAVVAGNSTTPRCAC TAGYHWS 122

QY 95 -DCALVNRQRANCS-----H---TSDAVCGDCLPGFY----- 12

Db 122 QDCECCR--NTECAPGLGAQHPLQLNKDVTCKPCLAGYFSDAFSSTDCKRPWTNCTFLG 179

QY 124 R K T K L V G F Q D M E C V P C G D P P --- P P Y E P H 149

Db 180 KREHHGTEKSDAVCSSLPARKPPNEPH 208

RESULT	12
W68288	
ID	W68288 standard; Protein; 616 AA.

PT New isolated receptor activator of necrosis factor-kappa B - useful
 PT for, e.g. developing products for regulating an immune or
 PT inflammatory response, treating toxic shock or sepsis
 PS Claim 27; Pages 48-51; 80pp; English.
 CC This represents the polypeptide RANK (receptor activator of necrosis
 CC factor-kappa (NF- κ B)). RANK is a member of the tumour necrosis factor
 CC (TNF) family. Host cells transformed or transfected with an expression
 CC vector comprising the RANK encoding nucleic acid can be used to produce
 CC recombinant RANK protein. The soluble RANK may be used for inhibiting
 CC activation of NF- κ B, by contacting a cell expressing membrane-associated
 CC RANK with a soluble RANK which binds to RANK ligand (RANKL). The soluble
 CC RANK polypeptide composition may also be used for regulating an immune
 CC or inflammatory response. Inhibition of NF- κ B by RANK antagonists may be
 CC useful in ameliorating negative effects of an inflammatory response that
 CC result from triggering of RANK, e.g. in treating toxic shock or sepsis,
 CC graft-versus-host reactions, or acute inflammatory reactions. They can
 CC also be used in adjunct therapy for disease characterised by neoplastic
 CC cells that express RANK. The products can also be used for detection and
 CC drug screening.

OY 124 -----RKTLYG--FQDMCVPCGDPPEPEH 149
DB 130 QDGDHCACRAATATSSPGQVOKGTESDTLCQNC--PPGTFFSPN 173

RESULT 15

W69238
ID W69238 standard; Protein: 283 AA.
AC W69238.
DT 21-OCT-1998 (first entry)
DE Herpesvirus entry mediator protein.
KW Entry mediator gene; herpesvirus; HVEM; tumour necrosis factor receptor;
KW gene expression regulator; cellular stress; inflammatory response;
KW lymphocyte activity regulator; autoimmune response.
OS Homo sapiens.
FN WO9825967-A1.
PD 18-JUN-1998.
PF 05-DEC-1997; U22278.
PR 12-DEC-1996; US-032705.
PA (GENE) GENENTECH INC.
PI Ashkenazi AJ, Marsters SA;
DR WPI: 98-348457/30.
N-PSDB: V44852.
PT Herpesvirus entry mediator polypeptide, HVEM - useful, e.g. in
PT assays for HVEM and to produce antibodies and transgenic animals,
PT e.g. for drug screening.
PS Claim 1: Fig 1: 46pp; English.
CC This sequence is the herpesvirus entry mediator (HVEM) protein of
CC the invention. The protein is useful in quantitative diagnostic assays
CC for HVEM, in affinity purification of HVEM from recombinant cells/natural
CC sources and in competitive-type receptor binding assays. It can also be
CC used to generate antibodies, also useful in diagnostic assays for HVEM
CC and affinity purification of HVEM. HVEM is believed to be a member of the
CC tumour necrosis factor receptor (TNFR) family, and transient transfection
CC of HVEM into human 293 cells caused marked activation of certain
CC transcription factors, e.g. AP-1, suggesting that HVEM is involved in
CC regulating gene expression in response to infectious stimuli and cellular
CC stress. The predominant expression of HVEM mRNA in lymphocyte-rich
CC tissues (e.g. spleen and peripheral blood) also suggested it may be a
CC receptor in regulating lymphocyte activity. Antibodies produced may
CC therefore be useful therapeutically, e.g. antagonistic antibodies may be
CC useful to block excessive inflammatory/autoimmune response resulting from
CC e.g. AP-1 induction. Whilst agonistic antibodies may enhance HVEM
CC regulation of such induction. The DNA may be used diagnostically, e.g.
CC to determine if DNA and/or RNA encoding HVEM is present in cells, and to
CC prepare HVEM polypeptide recombinantly. It is also useful to produce
CC non-human transgenic animals (e.g. mice or rats), especially knockout
CC animals containing cells with an altered gene encoding HVEM polypeptide.
CC Such animals are useful in the development and screening of
CC therapeutically useful reagents.
SQ Sequence 283 AA;

Query Match 14.7%; Score 128.5; DB 1: Length 283;
Best Local Similarity 27.7%; Pred. No. 0.00012;
Matches 46; Conservative 18; Mismatches 67; Indels 35; Gaps 10;

OY 10 RTVLFAILFLHLACKVSCETGDSRQOEFRDRSGNCVLCQKCGPMELSKECGFGYE 69
DB 17 RTDVLRLVLYLFLGAPCAPALPSCKEDEYVGS--EC--CPKCSGYRVKXKAGELTG- 73
OY 70 DAQCVPCRPFRKEDM--GFQCKPCADC--ALVNRFORANCSTSDAVCGDCLPGFY--- 124
DB 73 -TVCEPCPGIYIAHLNGSLKLCQCMCDPAMGLRASR--NCSRTENAVCG--CSPGHFCIV 129
OY 124 -----RKTLYG--FQDMCVPCGDPPEPEH 149
DB 130 QDGDHCACRAATATSSPGQVOKGTESDTLCQNC--PPGTFFSPN 173

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:04 ; Search time 33.52 Seconds

(without alignments)
168.750 Million cell updates/sec

Title: US-09-212-270-4

Perfect score: 873
Sequence: 1 MALKTPLHRTVLFALILFL.....QDMECVPCGDPPPPYEPHCE 151

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database: PIR_58:*

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	146.5	16.8	271	2	S12783	OX40 antigen precu
2	135.5	15.5	272	2	I48700	gene ox40 protein
3	126	14.4	1786	1	MMHUB1	laminin beta-1 cha
4	125	14.3	435	2	T54182	tumor necrosis fac
5	120	13.7	1786	1	MMMSB1	laminin beta-1 cha
6	118.5	13.6	277	2	A60771	B-cell activation
7	117	13.4	425	2	A26431	neure growth facto
8	117	13.4	349	2	D36858	gene G4R protein -
9	117	13.4	334	2	A48151	sperm tail protein
10	116	13.3	277	2	I37552	OX40 homologue - h
11	113	13.2	326	1	GOVZML	T2 protein - myxom
12	114	13.1	250	1	A49053	CD27 antigen precu
13	111.5	12.8	427	1	GOHUN	neure growth facto
14	111	12.7	289	2	A46515	B cell associated
15	111	12.7	305	2	A46476	CD40 - mouse
16	110.5	12.7	260	1	A46517	CD27 antigen precu
17	109	12.5	454	1	GOMST1	tumor necrosis fac
18	109	12.5	454	2	I57826	tumor necrosis fac
19	109	12.5	265	2	B48151	sperm tail protein
20	108.5	12.4	1607	1	MMMSB2	laminin gamma-1 ch
21	108.5	12.4	255	2	I38426	4-1BB - human
22	108.5	12.4	255	2	JT0752	lymphocyte activat
23	107.5	12.3	3084	1	MMMSA	laminin alpha-1 ch
24	106.5	12.2	1801	1	MMMSA	laminin beta-2 cha
25	106.5	12.2	3712	2	S18253	laminin alpha-1 ch
26	104.5	12.0	596	2	A45664	variant-specific s
27	104	11.9	459	2	I48854	gene murine tumour
28	103.5	11.9	416	2	JN0006	neure growth facto
29	103.5	11.9	1797	2	A55677	laminin beta-2 cha
30	103.5	11.9	2482	2	A47648	laminin beta-2 cha
31	103.5	11.9	2277	2	B47648	perlecan homolog u
32	103.5	11.9	1160	2	C47648	perlecan homolog u
33	103	11.8	1474	2	B38634	tumor necrosis fac
34	102.5	11.7	1548	2	S34583	serine protease
35	101.5	11.6	869	1	MMHUB2	laminin gamma-1 ch
36	101	11.6	869	1	JC4858	VLDL receptor prec
37	101	11.6	461	2	A35356	tumor necrosis fac
38	101	11.6	3075	2	S14458	laminin alpha-1 ch

40	100.5	11.5	1816	1	S68960	laminin alpha-4 ch
41	100.5	11.5	1394	2	A35626	transforming growt
42	99.5	11.4	873	1	QRRBV1	VLDL receptor prec
43	99.5	11.4	461	1	GQRRT1	tumor necrosis fac
44	99	11.3	256	2	B32393	T-cell antigen 4-1
45	98.5	11.3	873	1	A49729	VLDL receptor prec

ALIGNMENTS

RESULT 1
S12783
OX40 antigen precursor - rat
N:Alternate names: nerve growth factor receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
C:Accession: S12783; S08036
R:Allet, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A>Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphoc
A:Reference number: S12783; MWID:90214614
A:Accession: S12783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:g57830; PID:g57831
C:Keywords: growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-271/Product: OX40 antigen #status predicted <Mat>
F:211-235/Domain: transmembrane #status predicted <TM>

Query Match 16.8%; Score 146.5; DB 2; Length 271;
Best local Similarity 30.1%; Pred. No. 4.9e-06;
Matches 34; Conservative 14; Mismatches 50; Indels 15; Gaps 5;
QY 46 NCY-----LCKCGFMELEKCEGFEYGEDAOCVPCRPFRFEDMGFOCKPCADC 96
DB 25 NCVKDTPPSGKRCRCRCPOGHVSRSC--DHTDVTCHPCPEPFYNEAVNYDCKOCTQC 82
QY 97 ALVNRFO-RANGSHTSDAVCGDCLPGFYRRTKLVGSDMCEVCPCGPPPEYEP 148
DB 83 NHRSGSELKQNCPTDVTVC-QCRPGTQPRQDSSHLKGVDCVPC--PPGHFSP 132
RESULT 2
I48700
gene ox40 protein - mouse
N:Alternate names: OX40 antigen
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996
C:Accession: I48700; I48334; S34377
R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J
J. Immunol. 151, 5261-5271, 1993
A>Title: Cloning of mouse OX40: a T cell activation marker that may mediate T-B cell
A:Reference number: I48700; MWID:94044750
A:Accession: I48700
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-272 <RES>
A:Cross-references: EMBL:Z21674; NID:g312827; PID:g312828
R:Blitzland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
Eur. J. Immunol. 25, 926-930, 1995
A>Title: Gene structure and chromosomal localization of the mouse homologue of rat OX
A:Reference number: I48334; MWID:9555413
A:Accession: I48334
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-14, 'G', 16-272 <RE2>
A:Cross-references: EMBL:X85214; NID:g732818; PID:g732819
C:Genetics:
A:Gene: OX40
A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

RESULT 9
A48151
Sperm tail protein Mt598ca - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
C:Accession: A48151; MUID:93180818
R:Schaefer, M.; Borsch, D.; Hulster, A.; Schaefer, U.
MOL:Cell. Biol. 13, 1708-1718, 1993
A:Title: Expression of a gene duplication encoding conserved sperm tail proteins is tran
A:Reference number: A48151; MUID:93180818
A:Accession: A48151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1534 <SCH>
A:Cross-references: EMBL:X67704; NID:98267; PID:98268
A:Note: sequence extracted from NCBI backbone (NCBIP:125984)
C:Genetics:
A:Gene: Mt598ca
A:Cross-references: FlyBase:FBgn0002865

Query Match 13.4%; Score 117; DB 2; Length 334;
Best Local Similarity 29.4%; Pred. No. 0.0021;
Matches 40; Conservative 5; Mismatches 47; Indels 44; Gaps 9;

QY 34 CSROGEFDRSGNCVLCQCGFGM-ELSKEG-FGVEDAGCVPGRPHREKEDMGFKCK 91
DB 191 CFSTQYPAVEGRGCGPCGCGPCGRCPCGPG--PCGPGP-----RCG 237
QY 92 PCADCALVN-----RFQRANCSHTSDAVCGDCLPGFYRTKTLVGFQDMCVPCG 140
DB 238 PGCPAIVNCGCGCGTLMTSGFPVPAPCGPC--APCGPCP-----LCNSPCGPGC 285
QY 141 -----DPPPYE-PHC 150
DB 286 PCGCPSPCPCYESPEC 301

RESULT 10
137552
OX40 homologue - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: 137552
R:Latza, U.; Durkop, H.; Schmittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fonat
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homology: cDNA structure, expression and chromosomal assignment c
A:Reference number: 137552; MUID:94170844
A:Accession: 137552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:q472957; PID:q472958

Query Match 13.3%; Score 116; DB 2; Length 277;
Best Local Similarity 28.6%; Pred. No. 0.0022;
Matches 30; Conservative 14; Mismatches 43; Indels 18; Gaps 6;

QY 50 CKQCGPHELSKEGFGVEDAGCVPGRPHREKEDMGFKCKPCADCALVNFQRAN-CS 108
DB 43 CHECPGGMVSR--SRSONTVCRPCGCGFYNDVYSSKPCPCWCNCRSGSERKOLCT 100
QY 109 HTSDAVCGDCLPG-----FYRKTLLVGFQDMCVPCGDPPEYEP 148
DB 101 ATQDIIVC-RCRAGTQPLDSYK-----PGVDCAPC--PPGHFSP 135

RESULT 11
GOVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 24-Oct-1997

C:Accession: A40566
R:Upton, C.; Macen, J.L.; Schreiber, M.; Mcfadden, G.
Virology 184, 370-382, 1991
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosi
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Molecule type: DNA
A:Residues: 1-326 <UP>
A:Cross-references: GB:M95181; GB:M37976; NID:9332309; PID:9332310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:166,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.2%; Score 115; DB 1; Length 326;
Best Local Similarity 29.7%; Pred. No. 0.003;
Matches 35; Conservative 9; Mismatches 60; Indels 14; Gaps 6;

QY 13 LEFAILFLHLAC-----KVSCEFGDCSRQGEFDRSGNCVLCQCGPHELSKEGFG 66
DB 1 MRLTLLAVYACVYGGAIPYADGKCRGNRYED--GLC--CTSCPPGYSASRLC--G 54
QY 67 YGEDAGCVPGRPHREKEDMGF-QKCKPCADCALVNFQRANCSHTSDAVCGDCLPGFY 123
DB 55 PPSDIYVCSCKNETFTASTNHAPACVSCGRGCTGLHSESQCDKTRDRVC-DCSAGNY 111

RESULT 12
A49053
CD27 antigen precursor - mouse
N:Alternate names: CD27L receptor; T cell activation antigen CD27
C:Species: Mus musculus (house mouse)
C:Date: 19-Dec-1993 #sequence_revision 22-Apr-1995 #text_change 11-Sep-1998
C:Accession: A49053
R:Gravesteln, L.A.; Blom, B.; Nolten, L.A.; de Vries, E.; van der Horst, G.; Ossendor
Eur. J. Immunol. 23, 943-950, 1993
A:Title: Cloning and expression of murine CD27: comparison with 4-1BB, another lympho
A:Reference number: A49053; MUID:93209296
A:Accession: A49053
A:Molecule type: mRNA
A:Residues: 1-250 <GRA>
A:Note: sequence extracted from NCBI backbone (NCBIN:128168, NCBIP:128169)
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; homodimer; receptor; surface antigen; T-cell;
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-250/Product: CD27 antigen #status predicted <MAT>
F:21-182/Domain: extracellular #status predicted <EXT>
F:27-63/Domain: NGF receptor repeat homology <NG1>
F:65-105/Domain: NGF receptor repeat homology <NG2>
F:121-179/Region: proline/serine/threonine-rich
F:183-202/Domain: transmembrane #status predicted <TM>
F:203-250/Domain: intracellular #status predicted <INT>
F:95,162/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.1%; Score 114; DB 1; Length 250;
Best Local Similarity 28.1%; Pred. No. 0.003;
Matches 36; Conservative 15; Mismatches 41; Indels 36; Gaps 8;

QY 44 SGNVCYLCQCGPHELSKEGFGYED--AQCVCPR-HRREKEDMGFG-KKPCADCAL 98
DB 34 TGGGLCCRMCEGTFFVDC-----EQDRTAQCDCPICETISPPYHRRHESCRHCN- 89
QY 99 VNRQRANCSHTSDAVCGDCLPGFYRKTLLVGFQDMCVPCG-----D 141
DB 89 -SGFLIRNCTVYANAC-SC-----SKNQCRDQECTECDPINPALTRQPSRTPSPQ 139

QY 142 PPPPYEPH 149
DB 140 PPTTHPH 147

```

RESULT 13
GOHUN
Nerve growth factor receptor precursor, low affinity - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 01-May-1998
C:Accession: A25218; A60204; S21689; I57638
R:Johnson, D.; Lanahan, A.; Buck, C.R.; Segal, A.; Morgan, C.; Mercer, E.; Bothwell, M.
Cell 47, 545-554, 1986
A:Title: Expression and structure of the human NGF receptor.
A:Reference number: A25218; MID:87051725
A:Accession: A25218
A:Molecule type: mRNA
A:Residues: 1-427 <IOH>
A:Cross-references: GB:M14764; NID:g189204; PID:g189205
R:Marano, N.; Dietzschold, B.; Earley Jr., J.J.; Schatteman, G.; Thompson, S.; Grob, P.;
J. Neurochem. 48, 225-232, 1987
A:Title: Purification and amino terminal sequencing of human melanoma nerve growth factor.
A:Reference number: A60204; MID:87085574
A:Accession: A60204
A:Molecule type: protein
A:Residues: 29-31, 'T', '33-42, 'TT', '45-46, 'TX', '50-51, 'XX', '54-56 <MAR>
A:Experimental source: melanoma cell line A875
A:Note: this sequence has been corrected by a note added in proof to follow the nucleotide
R:Visavajhala, P.; Leszyk, J.D.; Lin-Goeke, J.; Ross, A.H.
Arch. Biochem. Biophys. 294, 244-252, 1992
A:Title: Structural domains of the extracellular domain of human nerve growth factor
A:Reference number: S21689; MID:92198017
A:Accession: S21689
A:Status: preliminary
A:Molecule type: protein
A:Residues: 183-208 <VIS>
R:Segal, A.; Patil, N.; Chao, M.
Mol. Cell. Biol. 8, 3160-3167, 1988
A:Title: A constitutive promoter directs expression of the nerve growth factor receptor
A:Reference number: I57638; MID:89096903
A:Accession: I57638
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RES>
A:Cross-references: GB:M21621; NID:g189206; PID:g189207
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
C:Comment: This protein is thought to form a high-affinity receptor when it associates
C:Comment: This receptor undergoes both N- and O-linked glycosylation.
C:Genetics:
A:Gene: GDB:NGFR
A:Cross-references: GDB:120234; OMIM:162010
A:Map position: 17q21-17q22
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor;
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-427/Product: nerve growth factor receptor #status experimental <MAT>
F:32-65/Domain: extracellular #status predicted <EXT>
F:32-65/Domain: NGF receptor repeat homology <NG1>
F:67-108/Domain: NGF receptor repeat homology <NG2>
F:109-147/Domain: NGF receptor repeat homology <NG3>
F:149-189/Domain: NGF receptor repeat homology <NG4>
F:197-248/Region: serine/threonine-rich
F:251-272/Domain: transmembrane #status predicted <TRM>
F:273-427/Domain: intracellular #status predicted <INT>
F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.8%; Score 111.5; DB 1; Length 427;
Best Local Similarity 30.4%; Pred. No. 0.0074;
Matches 35; Conservative 10; Mismatches 49; Indels 21; Gaps 7;

```

```

OY 72 QCVPCRPHREKED--WFOKRCPCADCALVNRFORANCSTSPAVCGDCLPGFR 124
DB 63 VCEPCLDSVTFPSDVVSATPEKRCFCTECVGLQSNM-APCYEADDAVC-RAKGYTQ 115

RESULT 14
B cell-associated surface molecule CD40 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockay
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; MID:93094586
A:Accession: A46515
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Experimental source: BALB/c, liver
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)

Query Match 12.7%; Score 111; DB 2; Length 289;
Best Local Similarity 27.1%; Pred. No. 0.0062;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

OY 13 LFAILLFLHLACKVCEFGDCSRQOEFDRSGNVCYLKQCGGMLSRGCGGYEDAQ 72
DB 10 LMGCLLTAHLG---OCVT--CSDKQYLHD--GQC--CDLQPGSRLTSHC--TALEKTQ 58

OY 73 CVPCRPHREKEDG-----FQCKRCPCADCALVNRFORANCST--SPAVC----- 116
DB 59 CHPCDSGERSAQNREIRCHQHRHCEP-----NGLRKKEGTAEPSDTVCTCKEGQHC 111

OY 116 -----GDCLPGFYRKTIVGFQDMCEVPC 139
DB 112 TSKDCENCAQHTPCIPGFGVMEKATETTTDVTCHPC 146

RESULT 15
CD40 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997
C:Accession: A46476
R:Torres, R.M.; Clark, E.A.
J. Immunol. 148, 620-626, 1992
A:Title: Differential increase of an alternatively polyadenylated mRNA species of mur
A:Reference number: A46476; MID:92105763
A:Accession: A46476
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-305 <TOR>
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)
C:Keywords: transmembrane protein

Query Match 12.7%; Score 111; DB 2; Length 305;
Best Local Similarity 27.1%; Pred. No. 0.0064;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

OY 13 LFAILLFLHLACKVCEFGDCSRQOEFDRSGNVCYLKQCGGMLSRGCGGYEDAQ 72
DB 10 LMGCLLTAHLG---OCVT--CSDKQYLHD--GQC--CDLQPGSRLTSHC--TALEKTQ 58

OY 73 CVPCRPHREKEDG-----FQCKRCPCADCALVNRFORANCST--SPAVC----- 116
DB 59 CHPCDSGERSAQNREIRCHQHRHCEP-----NGLRKKEGTAEPSDTVCTCKEGQHC 111

OY 116 -----GDCLPGFYRKTIVGFQDMCEVPC 139
DB 112 TSKDCENCAQHTPCIPGFGVMEKATETTTDVTCHPC 146

```

Thu May 20 13:49:56 1999

us-09-212-270-4.rpr

Page 7

Search completed: May 13, 1999, 20:59:04
Job time: 2318 sec

This Page Blank (uspio)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 1999, 14:43:38 ; Search time 25.8 Seconds
(without alignments)
157.089 Million cell updates/sec

Title: US-09-212-270-4
Perfect score: 873
Sequence: 1 MALKVPLHRTVLEFAILLFL.....QDMECVPCGDDPPPEPHCE 151

Scoring table: BLOSUM62

Searched: 74019 segs, 26840295 residues
Database: SwissProt_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being predicted,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	146.5	16.8	271	1 OX40_RAT	P15725 rattus norv
2	135.5	15.5	272	1 OX40_MOUSE	P47741 mus musculu
3	126	14.4	1786	1 LMB1_HUMAN	P07942 homo sapien
4	125	14.3	435	1 TNRC_HUMAN	P36941 homo sapien
5	123.5	14.1	417	1 WSL1_HUMAN	O93038 h wsl-1 pro
6	121	13.9	1799	1 LMB2_MOUSE	O61292 mus musculu
7	120	13.7	1786	1 LMB1_MOUSE	P02469 mus musculu
8	118.5	13.6	277	1 CD40_HUMAN	P25942 homo sapien
9	117	13.4	425	1 NGFR_RAT	P07174 rattus norv
10	117	13.4	349	1 VC22_VARY	P34013 variola vir
11	116	13.3	277	1 OX40_HUMAN	P43489 homo sapien
12	115	13.2	326	1 VT2_MYXVL	P28825 myxoma viru
13	114	13.1	250	1 CD27_MOUSE	P41272 mus musculu
14	111.5	12.8	427	1 NGFR_HUMAN	P08138 homo sapien
15	111	12.7	289	1 CD40_MOUSE	P27512 mus musculu
16	110.5	12.5	1798	1 LMB2_HUMAN	P55268 homo sapien
17	109.5	12.5	260	1 CD27_HUMAN	P26842 homo sapien
18	109	12.5	454	1 TNRI_MOUSE	P25118 mus musculu
19	108.5	12.4	255	1 41BB_MOUSE	O07011 homo sapien
20	108.5	12.4	1607	1 LMB1_MOUSE	P02468 mus musculu
21	107.5	12.3	3084	1 LMA1_MOUSE	P19137 mus musculu
22	106.5	12.2	3712	1 LMA_DROME	O00174 drosophila
23	106.5	12.2	1801	1 LMB2_RAT	P15800 rattus norv
24	106	12.1	3635	1 LMA5_MOUSE	O61001 mus musculu
25	106	12.1	415	1 TNRC_MOUSE	P50284 mus musculu
26	103.5	11.9	416	1 NGFR_CHICK	P18519 gallus galli
27	103.5	11.9	2481	1 UN52_CAEEL	O06561 caenorhabd1
28	103	11.8	474	1 TNR2_MOUSE	P25119 mus musculu
29	101.5	11.6	1609	1 LMG1_HUMAN	P11047 homo sapien
30	101	11.6	3075	1 LMA1_HUMAN	P25391 homo sapien
31	101	11.6	461	1 TNR2_HUMAN	P20303 homo sapien
32	100.5	11.5	1394	1 TGFH_HUMAN	P22064 homo sapien
33	100	11.5	1789	1 LMB1_DROME	P11046 drosophila
34	99.5	11.4	873	1 LDVR_RABIT	P35953 oryctolagus
35	99.5	11.4	461	1 TNRI_RAT	P22934 rattus norv
36	99	11.3	256	1 41BB_MOUSE	P20334 mus musculu
37	98.5	11.3	873	1 LDVR_HUMAN	P98135 homo sapien
38	98.5	11.2	1816	1 LMA4_HUMAN	O16363 homo sapien
39	97.5	11.2	873	1 LDVR_MOUSE	P98156 mus musculu
40	97	11.1	863	1 LDVR_CHICK	P98165 gallus galli
41	96.5	11.1	335	1 PASA_HUMAN	P25445 homo sapien
42	96	11.0	854	1 LDLR_CRICR	P33950 cricetus
43	95.5	10.9	612	1 UNC6_CAEEL	P34710 caenorhabd1

ALIGNMENTS

44 95.5 10.9 826 1 YTOJ_CAEEL Q19673 caenorhabd1
45 95 10.9 1680 1 FUR2_DROME P30432 drosophila

RESULT 1

ID	OX40_RAT	STANDARD:	PRT:	271 AA.
AC	P15725:			
DT	01-APR-1990 (REL. 14, CREATED)			
DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).			
GN	TXGP1L OR OX40.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EURARVOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-T-CELL.			
RX	MEDLINE: 90214614.			
RA	MALLET S., FOSSUM S., BARCLAY A.N.;			
RL	EMBO J. 9:1063-1068(1990).			
CC	- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	- TISSUE SPECIFICITY: ACTIVATED T-CELLS.			
CC	- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.			
DR	EMBL: X17037; G57831; -			
DR	PIR: S08036; S08036.			
DR	PIR: S12783; S12783.			
DR	PROSITE: PS00652; TNFR_NGFR.1; 3.			
DR	PROSITE: PS50050; TNFR_NGFR.2; 2.			
KW	RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;			
KW	SIGNAL.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	271	OX40L RECEPTOR.
FT	DOMAIN	20	210	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	211	235	POTENTIAL.
FT	DOMAIN	236	271	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	25	164	4 X TNFR-CYS.
FT	REPEAT	25	60	TNFR-CYS 1.
FT	REPEAT	61	102	TNFR-CYS 2.
FT	REPEAT	103	123	TNFR-CYS 3 (INCOMPLETE).
FT	REPEAT	124	164	TNFR-CYS 4.
FT	CARBOHYD	143	143	POTENTIAL.
SQ	SEQUENCE	271 AA:	29895 MW;	65460ED2 CRC32;
Query Match				
Best Local Similarity 16.8%; Score 146.5; DB 1; Length 271;				
Matches 34; Conservative 14; Mismatches 50; Indels 15; Gaps 5;				
OY	46 NCV-----LCKQCGRMELSKCGFGEDACQVPCRPFRHREKDWGFOKRCPCADC	96		
DB	25 NCVKTYPSGHKCRCPGHCVMGRS--DTRDRTVCHRCBEGFYNEAVNYDTCKQCTQC	82		
OY	97 ALVNFQ-RANSRSHSDAVCGCLPGFIRKTLVGFQDMECVPCGDDPPPIYP	148		
DB	83 NHRSSSELKUNCTPTEDTYVC-QCRPGTOPRODSSHKLGVDVPC--PGHFSP	132		
RESULT 2				
ID	OX40_MOUSE	STANDARD:	PRT:	272 AA.
AC	P47741:			
DT	01-FEB-1996 (REL. 33, CREATED)			
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).			
GN	TXGP1 OR OX40.			
OS	MUS MUSCULUS (MOUSE).			

```

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
  [1]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE: 94044750.
RA CALDERHEAD D.M., BUHLMANN J.E., VAN DEN BERTWEGH A.J.,
RA CLAASSEN E., NOELLE R.J., FELL H.;
RN J. IMMUNOL. 151:5261-5271(1993).
  [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95255413.
RA BIRELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA BARCLAY A.N.;
RN EUR. J. IMMUNOL. 25:926-930(1995).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: 221674; G312828; -.
DR MGI: X85214; G732819; -.
DR MGD: MGI:104512; TXGP1.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR RECEPTOR: T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT SIGNAL. 1 19 POTENTIAL.
FT CHAIN 20 272 OX40L RECEPTOR.
FT DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 236 POTENTIAL.
FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 165 4 X TNFR-CYS.
FT REPEAT 26 61 TNFR-CYS 1.
FT REPEAT 62 103 TNFR-CYS 2.
FT REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 125 165 TNFR-CYS 4.
FT CARBOHYD 144 144 POTENTIAL.
FT CONFLICT 15 15 A -> G (IN REF. 2).
SQ SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

```

Query Match 15.5%; Score 135.5; DB 1; Length 272;
 Best Local Similarity 29.4%; Pred. No. 1.8e-05;
 Matches 35; Conservative 14; Mismatches 43; Indels 27; Gaps 7;

```

OY 46 NCY-----LCKOCGPMELSKCEGFGYGDACVPCRPFRKEDMGFKCKACAD 96
  [1]
DB 26 NCKKHTYPSCHKRCRCQPHGMVSR--DHTDITLCPETGTFYNEAVNYDTCKCTOC 83
OY 97 ALVNRQ-RANCSHTSDAVCGDCLPGFYRTKLVGFD-----MCCVPCGDDPPPEP 148
  [1]
DB 84 NHRSGSELKONCTPDTCV-RCRPGTQPR-----QDSGYKLGVDCVPC--PPGHFSP 133

```

RESULT 3
 LMB1_HUMAN STANDARD; PRT; 1786 AA.

```

ID LMB1_HUMAN
AC P07942;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN LMB1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90368768.
RA VUOLTEENAHO R., CHOW L.T., TRYGGVASON K.;
RL J. BIOL. CHEM. 265:15611-15616(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87280097.

```

```

RA PIKKARAINEN T., EDDY R., FUKUSHIMA Y., BYERS M., SHOWS T.,
RA PIHLAJANIEMI T., SARASTE M., TRYGGVASON K.;
RN J. BIOL. CHEM. 262:10454-10462(1987).
  [3]
RP SEQUENCE OF 1276-1709 FROM N.A.
RX MEDLINE: 88021029.
RA JAYE M., MODI W.S., RICCA G.A., MUDD R., CHIU I.M., O'BRIEN S.J.,
RA DROHAN W.N.;
RN AM. J. HUM. GENET. 41:605-615(1987).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBUNIT: THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),
CC LAMININ-2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
DR EMBL: M61951; G186913; JOINED.
DR EMBL: M58147; G186913; JOINED.
DR EMBL: M61917; G186913; JOINED.
DR EMBL: M61918; G186913; JOINED.
DR EMBL: M61921; G186913; JOINED.
DR EMBL: M61922; G186913; JOINED.
DR EMBL: M61924; G186913; JOINED.
DR EMBL: M61925; G186913; JOINED.
DR EMBL: M61926; G186913; JOINED.
DR EMBL: M61927; G186913; JOINED.
DR EMBL: M61928; G186913; JOINED.
DR EMBL: M61929; G186913; JOINED.
DR EMBL: M61930; G186913; JOINED.
DR EMBL: M61931; G186913; JOINED.
DR EMBL: M61932; G186913; JOINED.
DR EMBL: M61933; G186913; JOINED.
DR EMBL: M61934; G186913; JOINED.
DR EMBL: M61935; G186913; JOINED.
DR EMBL: M61936; G186913; JOINED.
DR EMBL: M61938; G186913; JOINED.
DR EMBL: M61939; G186913; JOINED.
DR EMBL: M61940; G186913; JOINED.
DR EMBL: M61941; G186913; JOINED.
DR EMBL: M61942; G186913; JOINED.
DR EMBL: M61943; G186913; JOINED.
DR EMBL: M61944; G186913; JOINED.
DR EMBL: M61945; G186913; JOINED.
DR EMBL: M61946; G186913; JOINED.
DR EMBL: M61947; G186913; JOINED.
DR EMBL: M61948; G186913; JOINED.
DR EMBL: M61949; G186913; JOINED.
DR EMBL: M61950; G186913; JOINED.
DR EMBL: M55370; G186876; JOINED.
DR EMBL: M55378; G186876; JOINED.
DR EMBL: M55379; G186876; JOINED.
DR EMBL: M55385; G186876; JOINED.
DR EMBL: M55371; G186876; JOINED.
DR EMBL: M55372; G186876; JOINED.
DR EMBL: M55373; G186876; JOINED.
DR EMBL: M55374; G186876; JOINED.
DR EMBL: M55375; G186876; JOINED.
DR EMBL: M55376; G186876; JOINED.
DR EMBL: M55344; G186876; JOINED.
DR EMBL: M55345; G186876; JOINED.
DR EMBL: M55346; G186876; JOINED.

```


DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
 GN LTR OR TNECR.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC ETHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER.
 RX MEDLINE; 93252381.
 RA BAENS M., CHAFFANET M., CASSIMAN J.J., DEN BERGHE H., MARIYEN P.;
 RL GENOMICS 16:214-218(1993).
 RN [2]
 RP FUNCTION.
 RX MEDLINE; 94225209.
 RA CROME P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.,
 RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
 RL SCIENCE 264:707-710(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; L04270; G339762; -.
 DR HSSP; P19999; ICLG.
 DR MIM; 600979; -.
 DR PROSITE; PS00053; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 3.
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 FT SIGNAL 1 30
 FT CHAIN 1 30
 FT DOMAIN 31 435
 FT TRANSMEM 228 248
 FT DOMAIN 249 435
 FT DOMAIN 42 211
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 168
 FT REPEAT 169 211
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 62 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 148
 FT DISULFID 142 167
 FT DISULFID 170 185
 FT CARBOHYD 40 40
 FT CARBOHYD 177 177
 SQ SEQUENCE 435 AA; 46709 MW; 203882DD CRC32;

WSL1_HUMAN
 ID WSL1_HUMAN STANDARD; PRT; 417 AA.
 AC 093038; 093036; 093037; 092983; P78515; Q99831; Q99722; P78507;
 DT 01-NOV-1987 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE WSL-1 PROTEIN PRECURSOR (APOPTOSIS-MEDIATING RECEPTOR DR3) (APOPTOSIS-
 DE MEDIATING RECEPTOR TRAMP) (DEATH DOMAIN RECEPTOR 3) (WSL PROTEIN)
 DE (APOPTOSIS INDUCING RECEPTOR AIR) (APO-3) (LYMPHOCYTE ASSOCIATED
 DE RECEPTOR OF DEATH) (LARD).
 GN WSL1 OR WSL OR DR3 OR DDR3.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC ETHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS.
 RC TISSUE-LYMPHOID.
 RX MEDLINE; 97088617.
 RA KITSON J., RAVEN T., JIANG Y.-P., GOEDDEL D.V., GILES K.M., PUN K.-T.,
 RA GRINHAM C.J., BROWN R., FARROW S.N.;
 RL NATURE 384:372-375(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE; 97081063.
 RA CHINNATYAN A.M., O'ROURKE K., YU G.-L., LYONS R.H., GARG M.,
 RA DUAN D.R., XING L., GENTZ R., NI J., DIXIT V.M.;
 RL SCIENCE 274:990-992(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA DEGLI-ESPOSITO M.A., DIN W.S., COSMAN D., SMITH C.A., GOODWIN R.G.;
 RN SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HEART;
 RA MARSTERS S.A., SHERIDAN J.P., DONAHUE C.J., PITTI R.M., GRAY C.L.,
 RA GODDARD A.D., BAUER K.D., ASHKENAZI A.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA SKEATON G., XU X.N., OLSEN A., COMPER A., TAN R., MCMICHAEL A.,
 RA BELL J.;
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [6]
 RP SEQUENCE OF 4-417 FROM N.A.
 RC TISSUE-BRAIN, AND FETAL LUNG;
 RX MEDLINE; 97205335.
 RA BODMER J.-L., BURNS K., SCHNEIDER P., HOFMANN K., STEINER V.,
 RA THOME M., BORNAND T., HAHNE M., SCHROETER M., WILSON A., FRENCH L.E.,
 RA BROWNING J.L., MACDONALD H.R., TSCHOPP J.;
 RL IMMUNITY 6:79-88(1997).
 RN [7]
 RP SEQUENCE OF 7-417 FROM N.A.
 RC TISSUE-BRAIN;
 RA CHAUDHARY P.M., HOOD L.E.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: INDUCES APOPTOSIS AND ACTIVATES NUCLEAR FACTOR KAPPA-B
 CC (NF-KAPPA-B). DIRECTLY INTERACTS WITH TRADD ADAPTOR MOLECULE. MAY
 CC PLAY A ROLE IN REGULATING LYMPHOCYTE HOMEOSTASIS.
 CC -1- SUBUNIT: HOMODIMER. INTERACTS STRONGLY VIA THE DEATH DOMAINS WITH
 CC THE TNFR1-ASSOCIATED MOLECULE TRADD AND THE TNFR1 RECEPTOR TO
 CC ACTIVATE AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
 CC NF-KAPPA B SIGNALING.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMOCYTES AND
 CC LYMPHOCYTES. DETECTED IN LYMPHOCYTE-RICH TISSUES SUCH AS THYMUS,
 CC COLON, INTESTINE, AND SPLEEN. ALSO FOUND IN THE PROSTATE.
 CC -1- PTM: GLYCOSYLATED (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS WSL-1/LARD-1A (SHOWN HERE),
 CC WSL-S1/LARD-3, AND WSL-S2 ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL; Y09392; E281462; -.
 DR EMBL; Y09392; E281460; -.

```

DR EMBL: Y09392; E281461; -
DR EMBL: U72763; G1669512; -
DR EMBL: U83599; G1800297; -
DR EMBL: U83600; G1800299; -
DR EMBL: U78029; G1778764; -
DR EMBL: U74611; G1763293; -
DR EMBL: U94501; G2071949; -
DR EMBL: U94504; G2071955; -
DR EMBL: U73580; G1695825; -
DR EMBL: U73581; G1695927; -
DR EMBL: U83597; G1800293; -
DR PROSITE: PS00652; TNFR_NGFR.1; 2.
DR PROSITE: PS50050; TNFR_NGFR.2; 1.
DR PROSITE: PS50017; DEATH_DOMAIN.1.
DR RECEPTOR; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; SIGNAL;
KW REPEAT.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 417 WSL-1 PROTEIN.
FT DOMAIN 25 199 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 200 220 POTENTIAL.
FT DOMAIN 223 417 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 34 192 4 X TNFR-CYS.
FT REPEAT 34 71 TNFR-CYS 1.
FT REPEAT 72 115 TNFR-CYS 2.
FT REPEAT 116 163 TNFR-CYS 3.
FT REPEAT 164 192 TNFR-CYS 4.
FT DOMAIN 332 413 DEATH DOMAIN.
FT CARBOHYD 67 67 POTENTIAL.
FT CARBOHYD 106 106 POTENTIAL.
FT VARSPLIC 182 218 STIGSPERCAVCGRMFMVYVLLAGVPLTIGA ->
VIGBGRWPCGPRPANHPLDHPPLASQAPGICR (IN
WSL-SI/LARD-3).
FT VARSPLIC 219 417 MISSING (IN WSL-SI/LARD-3).
FT VARSPLIC 200 253 MEWVYLLAGLVPLLTATLTLYRHGMPKPLVTADGAG
MEALTPPATHTLS -> SRMGAGNARGGTGMDRGAGEG
NHPTPTSCFGCSGSCSLALMSPLCLAPP (IN WSL-
S2).
FT VARSPLIC 254 417 MISSING (IN WSL-S2).
FT MUTAGEN 354 354 L->A: SUPPRESSES HOMODIMERIZATION, TNFR1
INTERACTION, AND APOPTOSIS INDUCTION.
FT MUTAGEN 356 356 L->A: SUPPRESSES HOMODIMERIZATION, AND
TNFR1 INTERACTION.
FT MUTAGEN 373 373 D->A: SUPPRESSES HOMODIMERIZATION, AND
TNFR1 INTERACTION.
FT CONFLICT 4 6 RPR -> AAA (IN REF. 6).
FT CONFLICT 60 60 P -> H (IN REF. 7).
FT CONFLICT 167 167 P -> L (IN REF. 6 AND 7).
FT CONFLICT 312 312 A -> R (IN REF. 1).
FT CONFLICT 370 370 R -> L (IN REF. 1).
FT CONFLICT 381 381 R -> H (IN REF. 1).
FT SEQUENCE 417 AA: 45385 MW; 3AEFF0A5 CRC32;

Query Match 14.1%; Score 123.5; DB 1: Length 417;
Best Local Similarity 31.9%; Pred. No. 0.00029;
Matches 37; Conservative 12; Mismatches 34; Indels 33; Gaps 8;

```

```

DE LAMININ BETA-2 CHAIN PRECURSOR.
GN LAMB2 OR S-LAM OR LAMS.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129/J;
RX MEDLINE: 96278760.
RA DURKIN M.E., GAUTAM M., LOECHEL S., SANES J.R., MERLIE J.P.,
RA ALBRECHTSEN R., WEMER U.M.;
RL J. BIOL. CHEM. 271:13407-13416(1996).
RN [2]
RP SEQUENCE OF 348-428 FROM N.A.
RX TISSUE-LUNG;
RC MEDLINE: 94319092.
RA ABERDAM D., GALLIANO M.E., MATTEI M.-G., ORTONNE J.P., MENEGUZZI G.;
RL MAMM. GENOME 5:393-394(1994).
RN [3]
RP FUNCTION.
RC STRAIN-129/J;
RX MEDLINE: 95191650.
RA NOAKES P.G., GAUTAM M., MUDD J., SANES J.R., MERLIE J.P.;
RL NATURE 374:258-262(1995).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: LAMININ-3 (S-LAMININ) REGULATES THE FORMATION OF MOTOR
NERVE TERMINALS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBUNIT: THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ),
LAMININ-4 (S-MEROSIN), AND LAMININ-7 (NS-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: NEUROMUSCULAR SYNAPSE AND KIDNEY GLOMERULUS.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
DR EMBL: U43541; G1244720; -.
DR EMBL: U42624; G1244720; JOINED.
DR EMBL: X75928; G511042; -.
DR MGD; MGI:9916; LAMB2.
DR PROSITE: PS00022; EGF_1; 10.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ TYPE EGF; 12.
DR GLYCOPROTEIN; BASEMENT MEMBRANE; EXTRACELLULAR MATRIX; COILED COIL;
KW LAMININ EGF-LIKE DOMAIN; CELL ADHESION; REPEAT; SIGNAL.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1799 LAMININ BETA-2 CHAIN.
FT DOMAIN 36 283 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 284 555 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
V).
FT DOMAIN 286 349 LAMININ EGF-LIKE 1.
FT DOMAIN 350 412 LAMININ EGF-LIKE 2.
FT DOMAIN 413 472 LAMININ EGF-LIKE 3.
FT DOMAIN 473 524 LAMININ EGF-LIKE 4.
FT DOMAIN 525 555 LAMININ EGF-LIKE 5 (INCOMPLETE).
FT DOMAIN 556 782 LAMININ DOMAIN IV.
FT DOMAIN 783 1191 8 X LAMININ EGF-LIKE REPEATS (DOMAIN
III).
FT DOMAIN 784 831 LAMININ EGF-LIKE 6.
FT DOMAIN 832 877 LAMININ EGF-LIKE 7.
FT DOMAIN 878 927 LAMININ EGF-LIKE 8.
FT DOMAIN 928 986 LAMININ EGF-LIKE 9.
FT DOMAIN 987 1038 LAMININ EGF-LIKE 10.
FT DOMAIN 1039 1095 LAMININ EGF-LIKE 11.

```


FT DOMAIN 541 772 LAMININ DOMAIN IV.
 FT DOMAIN 773 1178 8 X LAMININ EGF-LIKE REPEATS (DOMAIN III).
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 DOMAIN II.
 FT DOMAIN 1398 1430 DOMAIN ALPHA.
 FT DOMAIN 1431 1786 DOMAIN I.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1368 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1448 1778 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 883 893 BY SIMILARITY.
 FT DISULFID 869 886 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.
 FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1179 INTERCHAIN (PROBABLE).
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1185 1785 INTERCHAIN (PROBABLE).
 FT CARBOHYD 120 120 POTENTIAL.
 FT CARBOHYD 356 356 POTENTIAL.
 FT CARBOHYD 519 519 POTENTIAL.
 FT CARBOHYD 677 677 POTENTIAL.
 FT CARBOHYD 1041 1041 POTENTIAL.
 FT CARBOHYD 1195 1195 POTENTIAL.
 FT CARBOHYD 1279 1279 POTENTIAL.
 FT CARBOHYD 1336 1336 POTENTIAL.
 FT CARBOHYD 1343 1343 POTENTIAL.

FT CARBOHYD 1487 1487 POTENTIAL.
 FT CARBOHYD 1533 1533 POTENTIAL.
 FT CARBOHYD 1542 1542 POTENTIAL.
 FT CARBOHYD 1643 1643 POTENTIAL.
 FT CONFLICT 1531 1534 POTENTIAL.
 FT CONFLICT 1749 1749 D -> N (IN REF. 2).
 SQ SEQUENCE 1786 AA; 196904 MW; 56CA7685 CRC32;
 Query Match 13.7%; Score 120; DB 1; Length 1786;
 Best Local Similarity 24.5%; Pred. No. 0.0018;
 Matches 45; Conservative 18; Mismatches 57; Indels 64; Gaps 11;
 OY 10 RTVEAALFLFLH---LACKVS-----CET--GDPSRQGEFKRSGNCVLCKQCPG 56
 DB 756 RNIIF-SISALIHQTGLCEDPQGSLSVCDPNPGGQCQCRNVVGRV-----CNRCAPG 809
 OY 57 MELSKGCGYGEADQACVPCRPNR-----FKEADGCF 87
 DB 810 T-----FGGPPNG-CKPCDCHLGGSASAFCDATITGCHQFQGIYARQCDRLPGYWG 861
 OY 88 QCKKPC-----ADCALVNRFORANCSTSDAVCGDCLPGFYRKTLYGFQDMECVPCGD 141
 DB 862 PSCQPCQCNHGLDDCDYTGECISCDYTTGHNCRCLAGYV-GDPIIGSGD-HCRPCPC 919
 OY 142 PPPP 145
 DB 920 PDGP 923
 RESULT 8
 CD40_HUMAN STANDARD; PRT; 277 AA.
 ID CD40_HUMAN
 AC P25942;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDM40).
 GN CD40.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89356608.
 RA STAMENKOVIC I., CLARK E.A., SEED B.;
 RL EMBO J. 8:1403-1410(1989).
 RN [2]
 RP 3D-STRUCTURE MODELLING OF 24-144.
 RX MEDLINE; 97189482.
 RA BAJORATH J., ARUFFO A.;
 RL PROTEINS 27:59-70(1997).
 RN [3]
 RP 3D-STRUCTURE MODELLING OF 26-186 IN COMPLEX WITH CD40L.
 RX MEDLINE; 98266353.
 RA SINGH J., GARBER E., VAN VLIJMEN H., KARPSUSAS M., HSU Y.-M.,
 RA ZHENG Z., NARSMITH J.H., THOMAS D.;
 RL PROTEIN SCI. 7:1124-1135(1998).
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC EMBL; X60592; E1175755; -.
 DR PIR; S04460; S04460.
 DR PDB; 1CDF; 01-APR-97.
 DR MIM; 109535; -
 DR PROSITE; PS00652; TNFR_NGFR_1; 1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 4.
 KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL;
 FT 3D-STRUCTURE.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 277 CD40L RECEPTOR.
 FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).

Query Match	Similarity	Score	DB 1;	Length	427;
Matches 35;	Conservative 10;	Mismatches 49;	Indels 21;	Gaps 7;	
17 ILFLHL-----ACVSGGTGDCSNQGEFKRSGNVCYLCNKGCGPMELSKGFGYGHDA 71	12.8%;	111.5;	DB 1;	Length 427;	
15 LILLILLGVSLGAGKAEKCPITG-----LYTHSGHC--CKACNIGEGVAYPOG---ANQT 62	30.4%;	Pred. No. 0.0033;			
72 QCVPRPRPRFKED--WGFQCKPCADCAALVNRFORANCSHTSDAVCGCDLPGFYR 124					
63 VCEPGLDSVTSIDVYVSATPECKPCIECGGLGMS--APCEVADDAVC--KCAIGYIQ 115					

Search completed: May 19, 1999, 14:43:39
Job time: 46 sec

ID CD40_MOUSE STANDARD; PRT: 289 AA.
AC P27512:
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
GN CD40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92105763.
RA TORRES R.M., CLARK E.A.;
RU J. IMMUNOL. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C;
RA TORRES R.M.;
RU SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=LIVER;
RX MEDLINE: 93094586.
RA GRIMALDI J.C., TORRES R., KOZAK C.A., CHANG R., CLARK E.A.,
RA HOWARD M., COCKAYNE D.A.;
RU J. IMMUNOL. 149:3921-3926(1992).
CC - FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M83312; G1553059; -
DR EMBL: M94126; G192526; -
DR EMBL: M94129; G192526; JOINED.
DR EMBL: M94128; G192526; JOINED.
DR EMBL: M94127; G192526; JOINED.
DR PIR: A46476; A46476.
DR HSP: P19438; 1TNR.
DR MGD: MGT.88336; CD40.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS50050; TNFR_NGFR_2; 4.
KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 289 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 289 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 POTENTIAL.
SQ SEQUENCE 289 AA; 32111 MW; D8D70A2C CRC32;

Query Match 12.7%; Score 111; DB 1; Length 289;
Best Local Similarity 27.1%; Pred. No. 0.0027;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

QY 13 LEAAILFLHLACVSCETGDSGKRSQOEKDSGNCVLCQKQSGMELSKEGFGYGEDAQ 72
DB 10 LMGCLLTVNHLG--QCVT--CSDKQYIHD--GQC--CDLQPSRLTSHC--TALEKTQ 58
QY 73 CVPCPRPRFKEDMG-----FOKCKPCADCALVNRFORANCST--SDAVC----- 116
DB 59 CHPCDSGSEFSQWNRREIRCHGRHCEP-----NQLRVKKEGTAESDTYCTCKEGQHC 111
QY 116 -----GDCLPGFYRRTKLVGFDMECVPC 139
DB 112 TSKDCACAGHTPCIPFGVMEATETTDVTCCHPC 146

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:58 ; Search time 48.52 Seconds

(without alignments)
171.693 Million cell updates/sec

Title:

US-09-212-270-4

Perfect score:

873

Sequence: 1 MALKVPLRHVLEFAILFL.....ODMECVFCGDPPEPPHCE 151

Scoring table:

BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_phage:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	146	16.7	267	4	002764	002764 oryctolagus
2	141.5	16.2	277	2	014866	014866 homo sapien
3	141.5	16.2	426	2	014865	014865 homo sapien
4	136	15.6	348	11	057277	057277 monkeypox v
5	136	15.6	349	11	057281	057281 monkeypox v
6	136	15.6	349	11	057099	057099 monkeypox v
7	136	15.6	349	11	057100	057100 monkeypox v
8	136	15.6	349	11	057101	057101 monkeypox v
9	136	15.6	349	11	057102	057102 monkeypox v
10	136	15.6	348	11	057108	057108 monkeypox v
11	132	15.1	348	11	057103	057103 monkeypox v
12	131	15.0	186	11	072735	072735 compox viru
13	130.5	14.9	350	11	057116	057116 compox viru
14	129	14.8	625	10	035305	035305 mus musculu
15	128.5	14.7	283	2	092956	092956 mus musculu
16	127	14.5	380	2	000280	000280 homo sapien
17	126	14.4	351	11	073559	073559 compox viru
18	125	14.3	372	2	000279	000279 homo sapien
19	124.5	14.3	355	11	085308	085308 compox viru
20	123.5	14.1	418	2	000275	000275 homo sapien
21	123.5	14.1	253	2	000276	000276 homo sapien
22	121	13.9	351	11	057117	057117 compox viru
23	120	13.7	350	11	057123	057123 compox viru
24	119	13.6	349	11	057284	057284 camelipox vi
25	119	13.6	349	11	057098	057098 camelipox vi
26	118.5	13.6	349	11	057305	057305 compox viru
27	117.5	13.5	348	11	057112	057112 variola vir
28	117.5	13.5	348	11	085407	085407 variola vir
29	117	13.4	334	3	024403	024403 drosophila

30	117	13.4	349	11	057110	057110 variola vir
31	117	13.4	349	11	057111	057111 variola vir
32	117	13.4	349	11	089118	089118 variola vir
33	117	13.4	349	11	089098	089098 variola vir
34	116	13.3	351	11	057121	057121 compox viru
35	113	12.9	401	10	070202	070202 mus musculu
36	113	12.9	349	11	057097	057097 camelipox vi
37	111	12.7	1737	2	075097	075097 homo sapien
38	110.5	12.7	326	11	057120	057120 compox viru
39	110.5	12.7	326	11	057122	057122 compox viru
40	110	12.6	347	11	057115	057115 compox viru
41	109	12.5	265	3	024404	024404 drosophila
42	109	12.5	349	11	057109	057109 variola vir
43	109	12.5	360	11	057118	057118 compox viru
44	109	12.5	347	11	057119	057119 compox viru
45	108	12.4	401	10	008712	008712 mus musculu

ALIGNMENTS

RESULT 1	
ID 002764	PRELIMINARY: PRT: 267 AA.
AC 002764	
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)	
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)	
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)	
DE OX40 PRECURSOR (FRAGMENT)	
OS ORYCTOLAGUS CUNICULUS (RABBIT)	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;	
OC LACOMORPHA; LEPORIDAE; ORYCTOLAGUS.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-CHB:HM;	
RL ISONO T, SETO A;	
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
DR EMBL: AB003911; D1020869; -	
DR PROSITE: PS00652; TNFR_NGFR_1; 2.	
DR PFM: PFM00020; TNFR_C6; 3.	
KW SIGNAL.	
FT NON_TER	1
FT SIGNAL	<1 18 POTENTIAL.
FT CHAIN	19 267 OX40.
SQ SEQUENCE	267 AA; 28489 MW; FB3C0C5E CRC32;
Query Match	16.7%; Score 146; DB 4; Length 267;
Best Local Similarity	34.7%; Pred. No. 2.7e-05;
Matches 33; Conservative 11; Mismatches 37; Indels 14; Gaps 6;	
QY 50 CKOCPGKELKECFGGEACVPCRPHEKEDMGFKCPGACALVNRFG---RA 105	
DB 37 CLECPGPGVMSRC--NSQDICHPCPEGFYNEVNTQACKPCTQC---NRRSGSEFQQ 91	
QY 106 NCSHTSDAVCGDCLPFGYKRTLVGFQ--DMECVPC 139	
DB 92 ECTHRTDYTC-RCRPG---TQPLNGYKKGVDCAFC 122	
RESULT 2	
ID 014866	PRELIMINARY: PRT: 277 AA.
AC 014866	
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)	
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)	
DE SOLUBLE DEATH RECEPTOR 3 BETA.	
GN DR3.	
OS HOMO SAPIENS (HUMAN).	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;	
OC CATARRHINI; HOMINIDAE; HOMO.	
RN [1]	

RP SEQUENCE FROM N.A.
 RA MARCOCHA K., RIBEIRO P., RENARD N., CHARLOT C., COIFFIER B.,
 RA SALES G.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF026071; G2570833;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR PFAM: PF00020; TNR_C6; 2.
 DR SEQUENCE 277 AA; 2911 MW; 06E165C7 CRC32;

Query Match 16.2%; Score 141.5; DB 2; Length 277;
 Best Local Similarity 33.1%; Pred. No. 6.5e-05;
 Matches 40; Conservative 12; Mismatches 36; Indels 33; Gaps 8;

QY 33 DCSRQGEFFKDRSG-----NC--VLCKQCG--PGMELSKGCGFGYGEDACQVPCRPFRFKE 83
 Db 88 ECARQACDEQASQVALENCSAVADTRCGCKPGMFV--EC-----QVSGCVSSSPF---- 137
 QY 84 DMGFOCKPCADCALVNRFORANCSTSDAVCGDCLPGFYRKTKLVGFQDMGCVPCGDDP 143
 Db 137 -----YCQPLDCGALHRRHRLCLSR-RDTDCGTCLPGFY-----HGDCVSCPTTP 183
 QY 144 P 144
 Db 184 P 184

RESULT 3
 014865 PRELIMINARY; PRT; 426 AA.

ID 014865
 AC 014865;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE DEATH RECEPTOR 3 BETA.
 DR3.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98113360.
 RA MARCOCHA K., RIBEIRO P., CHARLOT C., RENARD N., COIFFIER B.,
 RA SALES G.;
 RL "A new death receptor 3 isoform: expression in human lymphoid cell
 RT lines and non-Hodgkin's lymphomas.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 242:376-379(1998).
 DR EMBL: AF026070; G2570831;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR PFAM: PF00020; TNR_C6; 2.
 DR PFAM: PF00531; death; 1.
 DR SEQUENCE 426 AA; 45950 MW; 87E00821 CRC32;

Query Match 16.2%; Score 141.5; DB 2; Length 426;
 Best Local Similarity 33.1%; Pred. No. 9.4e-05;
 Matches 40; Conservative 12; Mismatches 36; Indels 33; Gaps 8;

QY 33 DCSRQGEFFKDRSG-----NC--VLCKQCG--PGMELSKGCGFGYGEDACQVPCRPFRFKE 83
 Db 88 ECARQACDEQASQVALENCSAVADTRCGCKPGMFV--EC-----QVSGCVSSSPF---- 137
 QY 84 DMGFOCKPCADCALVNRFORANCSTSDAVCGDCLPGFYRKTKLVGFQDMGCVPCGDDP 143
 Db 137 -----YCQPLDCGALHRRHRLCLSR-RDTDCGTCLPGFY-----HGDCVSCPTTP 183
 QY 144 P 144
 Db 184 P 184

RESULT 4
 057277

ID 057277 PRELIMINARY; PRT; 348 AA.

AC 057277;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 OC ORTHOPOXYVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1996(96-17), ZAIRE-1996 (96-16);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88543; G2738061;
 DR EMBL: U87841; G2738061;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR SEQUENCE 348 AA; 38212 MW; 1AAAB69 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 348;
 Best Local Similarity 26.3%; Pred. No. 0.00022;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RYVLFPAIILFLHLACKVSC-----ETGDCSRQGEFFKDRSGNVCYLCKQCGPG 56
 Db 2 RSYLYSYIILFL-----SCIINGRDIAHPAPSNGKC-KDNEYRSRN--LCCTSCPPG 50
 QY 57 MELSKGCGFGYGEDACQVPCRPFRF-KEDMGFOCKPCADCALVNRFORANCSTSDAVC 115
 Db 51 TYASRLC--DSKTNVCTCTCGSDFTSHNNHLDACLSCGRDSDVRRSNTTHNRIC 108
 QY 116 GDLPGFY-----RKTRL-VGF-----QDMGCVPCG 140
 Db 109 -ECSPGYCLLKGSSGCRCTCKTGKIGYGVSTSDVLCSPCG 154

RESULT 5
 057291 PRELIMINARY; PRT; 349 AA.

ID 057291
 AC 057291;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 OC ORTHOPOXYVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88144; G2738087;
 DR EMBL: U87842; G2738063;
 DR EMBL: U87994; G2738079;
 DR EMBL: U87995; G2738081;
 DR EMBL: U88143; G2738085;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR SEQUENCE 349 AA; 38295 MW; 7313FCF9 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 349;
 Best Local Similarity 26.3%; Pred. No. 0.00023;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RYVLFPAIILFLHLACKVSC-----ETGDCSRQGEFFKDRSGNVCYLCKQCGPG 56
 Db 2 RSYLYSYIILFL-----SCIINGRDIAHPAPSNGKC-KDNEYRSRN--LCCTSCPPG 50
 QY 57 MELSKGCGFGYGEDACQVPCRPFRF-KEDMGFOCKPCADCALVNRFORANCSTSDAVC 115

Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISKTCGIGVSGYSTGDIYCSPCG 154

RESULT 6

057099 PRELIMINARY; PRT; 349 AA.
 AC 057099;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SIERRA LEONE-1970 (70-0266);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87843; G2738065; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38321 MW; 5315315C CRC32;

Query Match

Best Local Similarity 15.6%; Score 136; DB 11; Length 349;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAHPHAPSNGKC-KDNEYRSRN--LCCLSCPG 50
 QY 57 MELSKCEGFGYGEDAQCVCPRPHRF-KEDMGFOKCKPCADCAVNFQORANCSTSDAVC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISKTCGIGVSGYSTGDIYCSPCG 154

RESULT 7

057100 PRELIMINARY; PRT; 349 AA.
 AC 057100;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIGERIA-1971 (71-0082);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87844; G2738067; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38239 MW; F8871DD2 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 349;
 Best Local Similarity 26.3%; Pred. No. 0.00023;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAHPHAPSNGKC-KDNEYRSRN--LCCLSCPG 50
 QY 57 MELSKCEGFGYGEDAQCVCPRPHRF-KEDMGFOKCKPCADCAVNFQORANCSTSDAVC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108

QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISKTCGIGVSGYSTGDIYCSPCG 154

RESULT 8

057101 PRELIMINARY; PRT; 349 AA.
 AC 057101;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1977 (77-0666);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87845; G2738069; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38311 MW; 4FAF8AF1 CRC32;

Query Match

Best Local Similarity 15.6%; Score 136; DB 11; Length 349;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAHPHAPSNGKC-KDNEYRSRN--LCCLSCPG 50
 QY 57 MELSKCEGFGYGEDAQCVCPRPHRF-KEDMGFOKCKPCADCAVNFQORANCSTSDAVC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISKTCGIGVSGYSTGDIYCSPCG 154

RESULT 9

057102 PRELIMINARY; PRT; 349 AA.
 AC 057102;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BENIN-1978 (78-3945);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87846; G2738071; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38308 MW; 17182545 CRC32;

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 17:21:19 ; Search time 42.79 Seconds

(without alignments)
96.896 Million cell updates/secTitle: US-09-212-270-5
Perfect score: 1061

Sequence: 1 MVOLTQGTLEQSLRREVSRLL.....RAKALNLSPTGTLGPKVL 205

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A_Geneseq_34.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1019	96.0	233	1 W37002	Human tumour necro
2	836	78.8	168	1 W37003	Human tumour necro
3	228.5	21.6	285	1 W38391	Homo sapiens neutri
4	229.5	21.6	285	1 W62461	Human T cell surfa
5	229.5	21.6	285	1 W73043	Tumour necrosis fa
6	227.5	21.4	285	1 W62460	Human T cell surfa
7	217	20.5	266	1 W62462	Human T cell surfa
8	197	18.6	150	1 W62463	Mouse T cell surfa
9	99.5	9.4	155	1 R44145	TNF RGD mutain #1.
10	98.5	9.3	161	1 P60656	Sequence of mature
11	98.5	9.3	161	1 R44151	TNF RGD mutain #7.
12	98.5	9.3	163	1 R44152	TNF RGD mutain #8.
13	97.5	9.2	145	1 P82045	Sequence of new tu
14	97.5	9.2	155	1 R44148	TNF RGD mutain #4.
15	96.5	9.1	155	1 R05174	Tumoricidal polype
16	96.5	9.1	157	1 R14893	Antitumoural polyp
17	96.5	9.1	216	1 W11702	Alpha-1-thymosin-a
18	96	9.0	235	1 R05328	Rat tumour necrosi
19	96	9.0	235	1 R05996	Rat tumour necrosi
20	95.5	9.0	169	1 P82606	Fusion protein der
21	95.5	9.0	157	1 R14892	Antitumoural polyp
22	95.5	9.0	155	1 R44146	TNF RGD mutain #2.
23	95.5	9.0	157	1 P62472	Tumour necrosis fa
24	94.5	8.9	158	1 P70635	Sequence of tumour
25	94.5	8.9	151	1 R11615	Tumour Necrosis fa
26	94.5	8.9	151	1 R11968	Tumour Necrosis fa
27	94.5	8.9	155	1 R22314	TNF polypeptide mu
28	94.5	8.9	652	1 W48650	Heavy chain of hma
29	93.5	8.8	171	1 P82607	Fusion protein der
30	93.5	8.8	176	1 P82608	Fusion protein der
31	93.5	8.8	150	1 P88151	Polypeptide with a
32	93.5	8.8	155	1 P61653	L-M cell prolifera
33	93.5	8.8	157	1 R14900	Antitumoural polyp
34	93.5	8.8	155	1 R22313	TNF polypeptide mu
35	93.5	8.8	150	1 R29809	TNF analogue #2. P
36	93.5	8.8	155	1 R44144	Wild type TNF. Tum
37	93.5	8.8	158	1 R44149	TNF RGD mutain #5.
38	93.5	8.8	158	1 R44150	TNF RGD mutain #6.
39	93.5	8.8	157	1 R62478	Tumour necrosis fa
40	93.5	8.8	157	1 R62479	Tumour necrosis fa
41	93.5	8.8	157	1 R63097	Tumour necrosis fa
42	93	8.8	154	1 R22312	TNF polypeptide mu
43	92.5	8.7	151	1 R13816	TNF analogue encod

ALIGNMENTS

44 92.5 8.7 151 1 R14305 TNF analogue encod
45 92.5 8.7 157 1 R20627 Synthetic hTNF mut

```
RESULT 1
W37002
ID W37002 standard; Protein: 233 AA.
AC W37002:
DE 14-APR-1998 (first entry)
DE Human tumour necrosis factor delta.
KM Human tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
KM tumour necrosis factor epsilon; T cell proliferation;
KM immune regulation; inflammatory response.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc_difference 231..233
FT Note="Not given in the specification but present
if the cDNA is fully decoded"
FT
FT
FT WO9733902-A1.
PD 18-SEP-1997.
PD 14-MAR-1996; U03774.
PD 14-MAR-1996; WO-003774.
PR (HUMA-) HUMAN GENOME SCI INC.
PA Gentz RL, Ni J, Yu G;
PI WPI: 97-470810/43.
DR N-PSDB: V00493.
DR
PT Nucleic acid encoding human tumour necrosis factors delta and
epsilon - useful for destroying cancer cells, or mediating T cell
proliferation, immune regulation and inflammatory responses etc.
PS Claim 18; Fig 1; 83pp; English.
CC The present sequence represents human tumour necrosis factor (TNF)
delta. The TNF can be used to treat patients deficient in this factor
(optionally by in vivo expression). It is a ligand for TNF and can
destroy (by inducing apoptosis) some transformed cell lines (for tumour
treatment); mediate cell activity and proliferation (including T cells
to stimulate an immune response to viral, bacterial or parasitic
infections, also to eliminate autoreactive T cells in e.g. type I
diabetes), and is functionally linked as a primary mediator of immune
regulation and the inflammatory response. A typical application of the
protein, its agonists or antagonists is prevention of septic shock,
inflammation, cerebral malaria, human immunodeficiency virus activation,
CC graft-host reaction, bone resorption (osteoporosis), rheumatoid
arthritis and cachexia, also to promote wound healing (by upregulating
cell adhesion) and to regulate hematopoiesis. The protein can also be
used to isolate its cognate receptors (or receptor-encoding genes).
CC The cDNA can be used to produce recombinant proteins, and fragments of
it to isolate full-length or related sequences, for diagnostic detection
CC of mutations (indicative of disease or susceptibility) and for
CC chromosome identification. Analysis of patient samples for presence of
CC the new protein (e.g. using specific antibodies) can be used
CC diagnostically.
SQ Sequence 233 AA:
```

```
Query Match 96.0%; Score 1019; DB 1; Length 233;
Best Local Similarity 97.1%; Pred. No. 1.7e-99;
Matches 199; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 1 MVOLTQGTLEQSLRREVSRLLQGTGSPONGEGYPMQSLPEQSSDALAEWENGRSRKRA 60
DB 29 MALNLTQGTLEQSLRREVSRLLQGTGSPONGEGYPMQSLPEQSSDALAEWENGRSRKRA 88
OY 61 VLTQKOKHSHVLTLPVPIATSKDSDVTEVWQPALRRGRLAQAGVRIQDAQVYLL 120
DB 89 VLSQKOKHSHVLTLPVPIATSKDSDVTEVWQPALRRGRLAQAGVRIQDAQVYLL 148
OY 121 YSOVLFQDVTFTMGVSVREGGROETLFRCLIRSMPSHPDRAYNSOYAGVYHLLQGITL 180
DB 149 YSOVLFQDVTFTMGVSVREGGROETLFRCLIRSMPSHPDRAYNSOYAGVYHLLQGITL 208
```

QY 181 SVIIPRAKLNLSPHGTFLGFVKL 205
 ||||||||||||||||||
 Db 209 SVIIPRAKLNLSPHGTFLGFVKL 233

RESULT 2
 W37003 standard; Protein: 168 AA.
 ID W37003
 AC W37003;
 DT 14-APR-1998 (first entry)
 DE Human tumour necrosis factor epsilon.
 KW Human; tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
 KM tumour necrosis factor epsilon; T cell proliferation;
 OS immune regulation; inflammatory response.
 PN Homo sapiens.
 MO9733902-A1.
 PD 18-SEP-1997.
 PF 14-MAR-1996; 003774.
 PR 14-MAR-1996; WO-003774.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Gentz RL, Ni J, Yu G;
 DR MPI: 97-470810/43.
 N-PSDB; V00494.
 PT Nucleic acid encoding human tumour necrosis factors delta and epsilon - useful for destroying cancer cells, or mediating T cell proliferation, immune regulation and inflammatory responses etc.
 PS Claim 20; Fig 2; 83pp; English.
 CC The present sequence represents human tumour necrosis factor (TNF) epsilon. The TNF can be used to treat patients deficient in this factor (optionally by in vivo expression). It is a ligand for TNF and can destroy (by inducing apoptosis) some transformed cell lines (for tumour treatment); mediate cell activity and proliferation (including T cells to stimulate an immune response to viral, bacterial or parasitic infections, also to eliminate autoreactive T cells in e.g. type I diabetes), and is functionally linked as a primary mediator of immune regulation and the inflammatory response. A typical application of the protein, its agonists or antagonists is prevention of septic shock, inflammation, cerebral malaria, human immunodeficiency virus activation, graft-host reaction, bone resorption (osteoporosis), rheumatoid arthritis and cachexia, also to promote wound healing (by upregulating cell adhesion) and to regulate haematopoiesis. The protein can also be used to isolate its cognate receptors (or receptor-encoding genes). The cDNA can be used to produce recombinant proteins, and fragments of it to isolate full-length or related sequences, for diagnostic detection of mutations (indicative of disease or susceptibility) and for chromosome identification. Analysis of patient samples for presence of the new protein (e.g. using specific antibodies) can be used diagnostically.
 CC Sequence 168 AA;

Query Match 78.8%; Score 836; DB 1; Length 168;
 Best Local Similarity 89.1%; Pred. No. 1,9e-80;
 Matches 164; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

QY 22 GTGSPONGEGYPMOSLPEQSSDALEWENGSRKRAVLTKOKKHSHLVHPINAT 81
 ||||||||||||||||||
 Db 1 GTGSPONGEGYPMOSLPEQSSDALEWENGSRKRAVLTKOKKHSHLVHPINAT 46
 QY 82 SKDDSDVTEVMOPALRRGGLAOGYGVRIODAGVYLLYSQVLFQDVFPTMGQVVSREG 141
 ||||||||||||||||||
 Db 46 -KNDSDVTEVMOPALRRGGLAOGYGVRIODAGVYLLYSQVLFQDVFPTMGQVVSREG 104
 QY 142 QGQOETLFRICRSMPSHPDAVNSQVSAQVPHLHQGDILSVIIPRAKLNLSPHGTFLG 201
 ||||||||||||||||||
 Db 105 QGQOETLFRICRSMPSHPDAVNSQVSAQVPHLHQGDILSVIIPRAKLNLSPHGTFLG 164

QY 202 FVKL 205
 ||||
 Db 165 FVKL 168
 RESULT 3

W58391
 ID W58391 standard; Protein: 285 AA.
 AC W58391;
 DT 11-SEP-1998 (first entry)
 DE Homo sapiens neutrokin alpha protein.
 KW neutrokin alpha; cell proliferation; differentiation; migration;
 KM cytotoxicity; cell death; treatment; tumour; infection; inflammation;
 KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
 KM fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
 OS cachexia; detection; diagnosis; drug screening.
 FH Homo sapiens.
 FT Key
 FT Domain
 FT Location/Qualifiers
 FT 1..46 "intracellular domain"
 FT /note="72"
 FT /note="transmembrane domain"
 FT 73..285
 FT /note="extracellular domain"
 FT Domain
 FT W09818921-A1.
 PN 07-MAY-1998.
 PD 25-OCT-1996; 017957.
 PF 25-OCT-1996; WO-017957.
 PR 25-OCT-1996; WO-017957.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Ni J, Yu G;
 DR MPI: 98-272216/24.
 N-PSDB; V30934.
 DR New isolated human Neutrokin alpha - used to develop products for diagnosis and treatment of e.g. tumours, infections,
 PT immunodeficiencies or autoimmune diseases
 PS Claim 17; Fig 1; 104pp; English.
 CC The sequence is that of the neutrokin alpha protein.
 CC Neutrokin alpha (NA) polypeptides modulate cell proliferation, differentiation, migration, cytotoxicity and cell death.
 CC They can be used to treat e.g. tumour and tumour metastasis, infections by bacteria, viruses and other parasites, immunodeficiencies, inflammatory diseases, lymphadenopathy, autoimmune diseases, graft versus host disease and to stimulate peripheral tolerance, destroy some transformed cell lines, mediate cell activation and proliferation, and are functionally linked as primary mediators of immune regulation and inflammatory responses. Such activity is useful for immune enhancement or suppression, myeloprotection, stem cell mobilisation, acute and chronic inflammatory control and treatment of leukaemia. They can also be used to stimulate wound healing and to treat fibrotic disorders including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They can also be used to regulate haematopoiesis, by regulating the activation and differentiation of various haematopoietic progenitor cells, e.g. to release mature leukocytes from the bone marrow following chemotherapy, and in stem cell mobilisation. NA may also be used to treat sepsis, NA antagonists can be used to prevent septic shock, inflammation, cerebral malaria, activation of the HIV virus, graft-host rejection, bone resorption, rheumatoid arthritis and cachexia (wasting or malnutrition). They can also be used to treat e.g. autoimmune diseases such as multiple sclerosis and insulin-dependent diabetes and inflammatory and infectious diseases such as silicosis, and sarcoidosis, idiopathic pulmonary fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxin shock, atherosclerosis, histamine-mediated allergic reactions and immunological disorders including late phase allergic reactions, chronic urticaria, and atopic dermatitis by inhibiting chemokine-induced mast cell and basophil degranulation and release of histamine. Ige-mediated allergic reactions such as allergic asthma, rhinitis and eczema, inflammatory pulmonary diseases, rheumatoid arthritis, inflammation, degenerative and inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome, subepithelial basement membrane fibrosis or adult respiratory distress syndrome. The products can also be used for detection, diagnosis and drug screening.
 CC Sequence 285 AA;

Query Match 21.6%; Score 229.5; DB 1; Length 285;
 Best Local Similarity 30.9%; Pred. No. 2e-16;
 Matches 67; Conservative 41; Mismatches 86; Indels 23; Gaps 7;

```

0Y      7 QTELDSLRREVSRLQCTGCPSONGECYPQSIPEOS--SDALEAME-----NGENSR 56
Db      73 QGDILASIRAELOGGHNAEKLPAA--GAGAPFACGELEAPAVTAGKIEEPAPDGCNNSONR 130
Oy      57 KRRARVLIOKKQHSHVLYHPINAT--SKDDSVTFEMMQPALRRGGIAOGCVAYIDA 115
Db      131 NKRAAIVGVEPIVTYODCLQLADSEFTIOLKSYTFEPMILSPFRSGALSEKENKILLVKET 190
Oy      116 GYLLYLISQVLFQDVFTMGQOVVSRE- ----GCGROETLFCRIKRSMP-SHPDRAYSQYS 168
Db      191 GFVFIFYGVGLYIDKTYAMGHILQKKVHVHFGDELSLVTLFRCIQNNPETLPN---NSCIS 247
Oy      169 AGVPHLHGCDILSVIIIPRARAKLNLSFHGTFFGFVKL 205
Db      248 AGIAKLTEEGDELDQLAIPREMAQISLDGDTVFEGALKL 284

RESULT          4
W62461         W62461 standard; Protein: 285 AA.
W62461:
Dr      05-OCT-1998   (first entry)
De      Human T cell surface antigen 63954 protein sequence #2.
Km      Human: 639554; primate; rodent; mouse; T cell surface antigen; mammal;
Kw      diagnosis; antigen-specific proliferation; cytokine production;
Kv      immune response; autoimmune disorder; rheumatoid arthritis;
Kx      systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
Os      Homo sapiens.
Pn      MO9827114-AAZ.
Pd      25-JUN-1998.
Pr      16-DEC-1997; U23321.
Pr      17-DEC-1996; US-033601.
Pa      (SCHE ) SCHERING CORP.
Pi      Gorman DM;
Dr      MPI: 98-362719/31.
Dr      N-P5DB: V39985.
Pr      New isolated polypeptide, 63954 - used to develop products for
Pr      treating e.g. autoimmune disorders, inflammation, tissue rejection,
Pr      cancer or degenerative conditions
Ps      Claim 1; Page 60-61; 69pp; English.
Cc      The present sequence is a human T cell surface antigen, designated
Cc      63954. The novel protein designated 63954 is expressed on T cells.
Cc      Protein 63954 can modulate antigen-specific proliferation and cytokine
Cc      production on effector cells and may potentiate immune cell expansion or
Cc      apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
Cc      molecule for regulation of T cell mediated cell activation, and may cause
Cc      a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
Cc      63954 can be used to modulate immune responses in abnormal situations,
Cc      e.g. autoimmune disorders, including Rheumatoid arthritis, systemic
Cc      lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
Cc      as acute and chronic inflammatory responses in which T cell activation,
Cc      expansion, and/or immunological T cell memory play an important role,
Cc      such as chronic inflammation or tissue rejection. The products can also
Cc      be used in the treatment of conditions associated with abnormal
Cc      physiology or development, including abnormal proliferation, e.g.
Cc      cancerous conditions, or degenerative conditions. The products can also
Cc      be used for detection, diagnosis and drug screening.
Se      Sequence 285 AA:

```

```

Query Match      21.6%: Score 229.5; DB 1; Length 285;
Best Local Similarity 30.9%: Pred. No. 2e-16;
Matches 67; Conservative 41; Mismatches 86; Indels 23; Gaps 7,
QY 7 QNELSLREVERSLQCTGSPONGBEYPMQSLPEQS--SDALEAME-----NKEBSR 56
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 73 QGDNLRLRELGCHHAKEKLPAA--GAAPAPAGAEADAVYAGAKIFEPAPAGEGNSNSNR 130
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 57 KRRAYLQKKQKHQSVLYLHLYPINAT--SKDSDVTEYMQPDLRRGGLAOGIVRIODA 115
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 131 NKRAVQGPPEETVYQDCLQILADSETPTIGKSTTFYPMILSERGALAEKENKILVKTET 190
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 116 GYVLLYSQVLFQDVTFTMGQVVSRE-----GGGRQETLFRCIKSRMP--SHPDVAVNSQFS 168

```

Db 191 GFFEFYGVGLTIDKTYAMGHIQRKRVHVEGDELSLTFLFRCLIONMPELPFN---NSCYS 247

RESULT 5

ID W73043 standard; Protein; 285 AA.

AC W73043;

DI 07-JAN-1999 (first entry)

DE Tumour necrosis factor homologue TL5 protein.

KM Tumour necrosis factor homologue TL5; vaccine; chronic;

KM acute inflammation; arthritis; septilemia; autoimmune disease;

KM inflammatory bowel disease; psoriasis; transplant rejection;

KM graft vs. host disease; infection; stroke; ischaemia;

KM acute respiratory disease syndrome; restenosis; brain injury; AIDS;

KM bone disease; cancer; lymphoproliferative disorder; atherosclerosis;

KM Alzheimer's disease.

OS Homo sapiens.

PN EP-869180-A1.

PD 07-OCT-1998.

PF 01-APR-1998; 302526.

PR 03-DEC-1997; US-984396.

PR 02-APR-1997; US-041797.

PA (SMK) SMITHKLINE BEECHAM CORP.

PI Hurlie MR, Young PR;

PI WPI; 98-508494/44.

DR N-PSDB; V58894.

PT New tumour necrosis factor homologue, TL5 - useful for diagnosis and treatment of Alzheimer's disease, AIDS and cancer

PS Claim 10; Page 18; 23pp; English.

CC The present sequence encodes a tumour necrosis factor homologue TL5 polypeptide sequence. TL5 polypeptides and antibodies are useful for identifying compounds which agonise and antagonise TL5, and these can be administered for treatment to inhibit TL5 activity (antagonist) or enhance TL5 activity (agonist). Gene therapy using the expression system can also be used to enhance TL5 activity. Diseases or susceptibility to a disease can be diagnosed by determining the presence or absence of a mutation in the TL5 protein. TL5 polynucleotides are useful for locating genes associated with disease by hybridisation to chromosomes. TL5 polypeptides and polynucleotides can be used, especially to raise an immune response (i.e. as vaccines) for the treatment of chronic and acute inflammation, arthritis, septilemia, autoimmune diseases (e.g. inflammatory bowel disease, psoriasis), transplant rejection, graft vs. host disease, infection, stroke, ischaemia, acute respiratory disease syndrome, restenosis, brain injury, AIDS, bone diseases, cancer (e.g. lymphoproliferative disorders), atherosclerosis, and Alzheimers disease.

SC Sequence 285 AA;

[illegible]

RESULT 6
ID W62460 standard; Protein: 285 AA.
AC W62460;
DT 05-OCT-1998 (first entry)
DE Human T cell surface antigen 63954 protein sequence #1.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;
diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Homo sapiens.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.
PI Gorman DM;
DR WPI: 98-362719/31.
DR N-PSDB: V39984.
PT New isolated polypeptide, 63954 - used to develop products for
treating e.g. autoimmune disorders, inflammation, tissue rejection,
cancer or degenerative conditions
PS Claim 1; Page 57-58; 69pp; English.
CC The present sequence is a human T cell surface antigen, designated
63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
production on effector cells and may potentiate immune cell expansion or
apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
molecule for regulation of T cell mediated cell activation, and may cause
a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
63954 can be used to modulate immune responses in abnormal situations,
e.g. autoimmune disorders, including rheumatoid arthritis, systemic
lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
as acute and chronic inflammatory responses in which T cell activation,
expansion, and/or immunological T cell memory play an important role,
such as chronic inflammation or tissue rejection. The products can also
be used in the treatment of conditions associated with abnormal
physiology or development, including abnormal proliferation, e.g.
cancerous conditions, or degenerative conditions. The products can also
be used for detection, diagnosis and drug screening.
SQ Sequence 285 AA:

Query Match 21.4%; Score 227.5; DB 1; Length 285;
Best Local Similarity 31.2%; Pred. No. 3.2e-16;
Matches 67; Conservative 40; Mismatches 85; Indels 23; Gaps 7;

QY 9 ELQSLREVSRIGTGSGPQNGEGYPMQSLPEQS--SDALEWE-----NGESRRK 58
DB 75 DLASLRARELQGHAEKLP--GAGAPKAGLEAPAYTAGKITFEPPAPGEGNSQNSRKN 132
QY 59 RAVLTQKOKKHSHVHLVPIINAT--SKDDSDVTEVMQPALRGRLQAQGYRIODAGV 117
DB 133 RAVQGEFEFTVQOCLOLIADSETPTIQKSYTFVPMILSKRGSALAEKKNKILVETG 192
QY 118 YLXISQVLFQDVTFTMGVYSRE-----GQROETLFRCTRSM--SHPDRAVNSQYSA 170
DB 193 FFLYGGVLYTKTYAMGHLIQKRVHVFGEDELSTVTLFRCIQNMPELTPN--NSCYSAG 249
QY 171 VPHLHQGDILSVIIPRAKALNLSPHGTFGLFYKL 205
DB 250 IAKLEGGDELQALIPRENAQISLDGDTVFFGALKL 284

RESULT 7
ID W62462 standard; Protein: 266 AA.
AC W62462;
DT 05-OCT-1998 (first entry)
DE Human T cell surface antigen 63954 protein sequence #3.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;

KW diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Homo sapiens.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.
PI Gorman DM;
DR WPI: 98-362719/31.
DR N-PSDB: V39986.
PT New isolated polypeptide, 63954 - used to develop products for
treating e.g. autoimmune disorders, inflammation, tissue rejection,
cancer or degenerative conditions
PS Claim 1; Page 63-64; 69pp; English.
CC The present sequence is a human T cell surface antigen, designated
63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
production on effector cells and may potentiate immune cell expansion or
apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
molecule for regulation of T cell mediated cell activation, and may cause
a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
63954 can be used to modulate immune responses in abnormal situations,
e.g. autoimmune disorders, including rheumatoid arthritis, systemic
lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
as acute and chronic inflammatory responses in which T cell activation,
expansion, and/or immunological T cell memory play an important role,
such as chronic inflammation or tissue rejection. The products can also
be used in the treatment of conditions associated with abnormal
physiology or development, including abnormal proliferation, e.g.
cancerous conditions, or degenerative conditions. The products can also
be used for detection, diagnosis and drug screening.
SQ Sequence 266 AA:

Query Match 20.5%; Score 217; DB 1; Length 266;
Best Local Similarity 29.6%; Pred. No. 3.7e-15;
Matches 64; Conservative 38; Mismatches 74; Indels 40; Gaps 7;

QY 7 QTELASREVSRIGTGSGPQNGEGYPMQSLPEQS--SDALEWE-----NGESRRK 56
DB 73 QGDLSLRARELQGHAEKLP--GAGAPKAGLEAPAYTAGKITFEPPAPGEGNSQNSRKN 130
QY 57 KRAVLTQKOKKHSHVHLVPIINATSKDDSDVTEVMQPALRGRLQAQGYRIODAGV 116
DB 131 NKRAV-----GQPEETGSYTFVPMILSKRGSALAEKKNKILVETG 172
QY 117 YLXISQVLFQDVTFTMGVYSRE-----GQROETLFRCTRSM--SHPDRAVNSQYSA 169
DB 173 YFFLYGGVLYTKTYAMGHLIQKRVHVFGEDELSTVTLFRCIQNMPELTPN--NSCYSAG 229
QY 170 GVPHLHQGDILSVIIPRAKALNLSPHGTFGLFYKL 205
DB 230 GIAKLEGGDELQALIPRENAQISLDGDTVFFGALKL 265

RESULT 8
ID W62463 standard; Protein: 150 AA.
AC W62463;
DT 05-OCT-1998 (first entry)
DE Mouse T cell surface antigen 63954 protein sequence.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;
diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Mus sp.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.

PI Gorman DM;
DR WPI: 98-362719/31.
DR N-PSDB: V39987.
PT New isolated polypeptide, 63954 - used to develop products for
PT treating e.g. autoimmune disorders, inflammation, tissue rejection,
PT cancer or degenerative conditions
PS Claim 1; Page 66; 69pp; English.
CC The present sequence is a mouse T cell surface antigen, designated
CC 63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
CC production on effector cells and may potentiate immune cell expansion or
CC apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
CC molecule for regulation of T cell mediated cell activation, and may cause
CC a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
CC 63954 can be used to modulate immune responses in abnormal situations,
CC e.g. autoimmune disorders, including rheumatoid arthritis, systemic
CC lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
CC as acute and chronic inflammatory responses in which T cell activation,
CC expansion, and/or immunological T cell memory play an important role,
CC such as chronic inflammation or tissue rejection. The products can also
CC be used in the treatment of conditions associated with abnormal
CC physiology or development, including abnormal proliferation, e.g.
CC cancerous conditions, or degenerative conditions. The products can also
CC be used for detection, diagnosis and drug screening.
SQ Sequence 150 AA;

Query Match 18.6%; Score 197; DB 1; Length 150;
Best Local Similarity 36.3%; Pred. No. 2.2e-13;
Matches 45; Conservative 26; Mismatches 43; Indels 10; Gaps 3;

OY 89 TEVMMQPALRRRGQAGQGYVRIDAGYLLYSQVLFQDVFTFMQVYSRE-----GQ 142
DB 29 TFEVPLSKRGNALAEKKNKTVROTGFYISQVLYTDPFPMCHVGRKRVHVGDE 88
OY 143 GHOETLFRCIRSRP-SHPRAYNSOYACVPHLHOGDILSVIPRAKLNLSPHGTFPG 201
DB 89 LSLVTLFRICIQMFKTLPN---NSCYSAGIARLEDEDEIQLAIPREMAQISRNQDITFG 145
OY 202 FVKL 205
DB 146 ALKL 149

RESULT 9
R44145
ID R44145 standard; Protein: 155 AA.
AC R44145;
DE 18-MAY-1994 (first entry)
DE TNF RGD muteln #1.
KW Wild type; tumour necrosis factor; RGD motif; human; TNF; muteln;
KW antitumour activity; cancer; metastasis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 56 /note= "Encoded by ATG"
FT J05271289-A.
PD 19-OCT-1993.
PE 17-MAR-1992; 108342.
PR 17-MAR-1992; JP-108342.
PA (ISHH) ISHIMURA SANGYO KAISHA LTD.
DR WPI: 93-362343/46.
PT Tumour necrosis factor muteln sequence - and recombinant plasmid
PT contg. DNA coding for it, and transformed microorganisms used for
PT its prodn.
PS Claim 2; Page 37; 56pp; English.
CC The sequences given in R44145-52 represents tumour necrosis factor
CC (TNF) mutelns in which amino acids 1-8 of wild type TNF are substituted
CC by an amino acid sequence comprising 3-16 amino acids and containing
CC at least one Arg-Gly-Asp motif. These human TNF mutelns have the same
CC antitumour activity as wild type TNF and show no cancer metastasis
CC promoting activity. These mutelns can be produced by recombinant DNA
CC techniques and may be used in drug compositions.

SQ Sequence 155 AA;

Query Match 9.4%; Score 99.5; DB 1; Length 155;
Best Local Similarity 30.1%; Pred. No. 0.0039;
Matches 41; Conservative 18; Mismatches 48; Indels 29; Gaps 7;

OY 80 ANSKDSD--VTEVMMQPAL-----RRGRLAQAGYGV-----IQDAGYLLYSQV 124
DB 1 SSSRGDSKPVARHAVVNPQAEQOLQNLNRRANLLANGVELRNQLVPSSEGILYISQV 60
OY 125 LFPD-----VFETMGQV-VSRGQGRQETLFR--CIRSMPSHPDRA--YNSQYSAG 170
DB 61 LKRGQCSPTHTVLTHTMRINAVSYOTKYNLSAIRSPQRETPDEGAEPWEPYILGG 120
OY 171 VPHLHOGDILSVIIPR 186
DB 121 VFQLEKGRDLSAQINR 136

RESULT 10
P60656
ID P60656 standard; Protein: 158 AA.
AC P60656;
DE 28-JUL-1991 (first entry)
DE Sequence of mature human tumour necrosis factor (hTNF) muteln
DE Ser 69 in PAW/31.
KW Antitumour; anticancer.
OS Homo sapiens.
PN WO8604606-A.
PD 14-AUG-1986.
PE 03-FEB-1986; U00236.
PR 19-OCT-1982; US-435154.
PR 15-APR-1983; US-486162.
PR 20-DEC-1983; US-564224.
PR 15-OCT-1984; US-661026.
PR 07-FEB-1985; US-698939.
PR 07-FEB-1985; US-695934.
PA (CETU) CETUS CORP.
PI Mark DF, Lin LS, Lu SDY, Wang AM;
DR WPI: 86-225458/34.
DR N-PSDB: N60558.
PT New synthetic mutelns of human tumour necrosis factor protein -
PT are Obtd. by direct mutagenesis and retain antitumour activity
PS Disclosure; Fig 3a; 47pp; English.
CC The sequence encoding TNF produced by the promyelocytic leukemia
CC cell line (HL-60, ATCC no CCL240) has been cloned and expressed in
CC E.coli (see N60557). Neither of the cysteine residues (69 and 101) in
CC the TNF sequence appears to be involved in disulphide linkages. The
CC patentors claim a novel synthetic muteln of a biologically active
CC hTNF protein, having at least one cysteine residue free from a
CC disulphide link and non-essential to the activity and having at
CC least one of the cysteine residues deleted or replaced by another AA.
CC Plasmid pAM731 (Ser 69) is claimed.
SQ Sequence 158 AA;

Query Match 9.3%; Score 98.5; DB 1; Length 158;
Best Local Similarity 28.1%; Pred. No. 0.0031;
Matches 39; Conservative 20; Mismatches 53; Indels 27; Gaps 6;

OY 75 LVPINATSKDSDVTEVMMQPAL-----RRGRLAQAGYGV-----IQDAGYLLYSQV 121
DB 1 MVRSSSRTPSDKPVARHAVVNPQAEQOLQNLNRRANLLANGVELRNQLVPSSEGILYI 60
OY 122 SQVLFQD-----VFETMGQV-VSRGQGRQETLFR--CIRSMPSHPDRA--YNSQ 167
DB 61 SQVLFQDQSPSTHTVLTHTMRINAVSYOTKYNLSAIRSPQRETPDEGAEPWEPYI 120
OY 168 SAVGPHLHOGDILSVIIPR 186
DB 121 LGGVFOLEKGRDLSAQINR 139

EH	Key	Location/Qualifiers
ET	misc_difference 59	
ET	/note= "Encoded by ATG"	
PN	J05271289-A.	
PD	19-OCT-1993.	
PF	17-MAR-1992; 108342.	
PR	17-MAR-1992; JP-108342.	
PA	(ISHH) ISHIHARA SANGYO KAISHA LTD.	
DR	WPI: 93-365243/46.	
PT	Tumour necrosis factor mutein sequence - and recombinant plasmid	
PT	contg. DNA coding for it, and transformed microorganisms used for	
PT	its prodn.	
CS	Claim 2; Page 38; 56pp; English.	
CC	The sequences given in R44145-52 represents tumour necrosis factor	
CC	(TNF) mutants in which amino acids 1-8 of wild type TNF are substituted	
CC	by an amino acid sequence comprising 3-16 amino acids and containing	
CC	at least one Arg-Gly-Asp motif. These human TNF muteins have the same	
CC	antitumour activity as wild type TNF and show no cancer metastasis	
CC	promoting activity. These muteins can be produced by recombinant DNA	
CC	techniques and may be used in drug compositions.	
CC	Sequence 158 AA;	
50		

Query Match	9.2%;	Score 97.5;	DB 1;	Length 158;
Best Local Similarity	30.4%;	Pred. No. 0.0066;		
Matches	41;	Conservative	16;	Mismatches 49;
			Indels	29;
			Gaps	7;

QY	81	TSKDDSD--VYEVAMQPL-----RRGGLDAAQYGR-----IQDDAGVILYSQVL	125
Db	5	TPRGSDSPKFAHVAVANPQAEQQLQMLNRRANALLANGVELRDNLQVYVPSSEGVILYISQVL	64
QY	126	FQD-----VYFTMQQV--VSRSGGGRQETLER--CISMSPHPRA--YNSQISAGV	171
Db	65	FKGGQSPETHVLLHHTHSRIANVSTQTKVNLSTAIKSPQARETPGGAERAKPWEIETVIGV	124
QY	172	PHLHGGDILSVIIPR	186
Db	125	FQLEKGGRLSAQINR	139

RESULT 15

R05174

ID R05174 standard; protein; 155 AA.

AC R05174;

DT 18-JUL-1990 (first entry)

DE Tumoricidal polypeptide.

KW Tumoricidal; Tumour necrosis factor.

PN J02019398-A.

PD 23-JAN-1990.

PF 06-JUL-1988; 156913.

PR 06-JUL-1988; JP-166913.

PA (ASAHI) Asahi Chemical Ind KK.

DR WPI: 90-064642/09.

DR N-PSDB: Q02034.

PT Tumoricidal polypeptide(s) - in which DNA to code the polypeptide(s) is constructed from human-derived TNF-coding DNA and rabbit-derived TNF-coding DNA.

PS Disclosure; Page 3; 11pp; Japanese.

CC The sequence is derived from tumour necrosis factor DNA.

CC See also R05173-76 and R05189-90.

Sequence 155 AA;

50

Query Match	9.1%;	Score 96.5;	DB 1;	Length 155;
Best Local Similarity	27.7%;	Pred. No. 0.0081;		
Matches	44;	Conservative	24;	Mismatches 54;
			Indels	37;
			Gaps	10;

```

0Y 79 NATSADSDYEVNMMQPAL-----RRGGLOAGCYVR-----IQDAGVILYSQVL 125
   :::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 2 SSRTPSPDKPAHVAVVANPAEGCOLMNRANALLANGVELELDNOLVYPSRGILYISQVL 61
   :::::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 126 EQD-----VTFITMGQV-VSRREGCGROETLECRIRSMPSH---PDRA-----YNSQY 167
   |:::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

D6 62 FKGGCSTHTLLHTISRLASVQ---TKVNLISLAKS-PCHRETPEEAEAPNAMYPIY 117

Q7 168 SAGVPHLHQGDILSVIIPRAKAKNLSPHG-TFLGFVKL 205
11 1 1 1 1 1 : : 1 1 : : 1 : 1
118 LGVAFQLEKGRDRLSTEVNQPE-YIDLAESQVYFGITALL 155

Search completed: May 13, 1999, 17:21:20
Job time: 20719 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:04 ; Search time 33.52 Seconds

(Without alignments)
229.098 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061
Sequence: 1 MWOLNQQTFLQSLRRVSRLL.....RARAKLNSPHGTFLEGVKL 205

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database: PIR.58.*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	96	9.0	235	2	JU0029	tumor necrosis fac
2	92	8.7	233	2	S22052	tumor necrosis fac
3	91.5	8.6	233	1	OMHUN	tumor necrosis fac
4	88	8.3	185	2	S52715	tumor necrosis fac
5	87.5	8.2	235	1	QWMSN	tumor necrosis fac
6	87	8.2	233	2	S24642	tumor necrosis fac
7	86.5	8.2	233	2	S11688	tumor necrosis fac
8	86	8.1	197	2	JH0309	tumor necrosis fac
9	85.5	8.1	205	1	OMHUX	lymphotoxin alpha
10	85.5	8.1	234	2	Q01344	tumor necrosis fac
11	85.5	8.1	234	2	A25451	tumor necrosis fac
12	82.5	7.8	235	2	S15490	tumor necrosis fac
13	81	7.6	392	2	I54491	cell surface antiq
14	80.5	7.6	357	2	D71262	probable hemolysin
15	80	7.5	234	2	JH0529	tumor necrosis fac
16	80	7.5	233	2	S13114	tumor necrosis fac
17	80	7.5	707	2	S78538	site-specific reco
18	78	7.4	232	2	S12606	tumor necrosis fac
19	78	7.4	281	2	I38707	Fas ligand - human
20	77	7.3	505	2	I37206	protein-tyrosine k
21	77	7.3	193	2	S06192	tumor necrosis fac
22	76	7.2	171	2	B44126	ribosomal protein
23	75.5	7.1	484	2	S48403	sc/svp protein - yea
24	75.5	7.1	753	2	S18689	phosphoribosylg
25	74	7.0	393	2	E64485	pleiotropic drug r
26	74	7.0	1068	2	S60855	hypothetical prote
27	74	7.0	486	2	S00855	tumor necrosis fac
28	73.5	6.9	202	2	B27303	nitrate transpor
29	73	6.9	670	2	S71787	probable TMV resis
30	73	6.9	1996	2	F71405	translation elonga
31	72	6.8	726	1	FC6430	translational relea
32	72	6.8	360	1	FC6430	translational relea
33	72	6.8	1586	2	S39580	HRM protein - hum
34	72	6.8	1572	2	S45251	SNF2alpha protein
35	72	6.8	749	2	S64780	hypothetical prote
36	72	6.8	1120	2	S67208	hypothetical prote
37	71.5	6.7	513	1	RGECAV	transcription regu
38	71.5	6.7	775	2	B70449	phenylalanyl-tRNA
39	71.5	6.7	337	2	S21405	porin - Neisseria

40	71.5	6.7	608	2	A64992	sensor protein Ato
41	71	6.7	416	2	JN0006	nerve growth facto
42	71	6.7	531035	2	S31035	retrovirus-related
43	71	6.7	717	2	S31034	retrovirus-related
44	71	6.7	179	2	B64861	hypothetical prote
45	71	6.7	638	2	S67605	hypothetical prote

ALIGNMENTS

RESULT 1

JU0029
tumor necrosis factor alpha precursor - rat
N/Alternate names: cachectin; TNF alpha
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Mar-1998
R/Shirai, T.; Shimizu, N.; Horiguchi, S.; Ito, H.
Agric. Biol. Chem. 53, 1733-1736, 1989
A/Title: Cloning and expression in Escherichia coli of the gene for rat tumor necrosi
A/Reference number: JU0029
A/Accession number: JU0029
A/Molecule type: DNA
A/Residues: 1-235 <SHI>
R/Kwon, J.; Chung, I.Y.; Benveniste, E.N.
Gene 132, 227-236, 1993
A/Title: Cloning and sequence analysis of the rat tumor necrosis factor-encoding gene
A/Reference number: JN0868
A/Accession number: JN0868
A/Molecule type: DNA
A/Residues: 1-235 <KWO>
A/Cross-references: GB:L00981; NID:g205253; PID:g205254
R/Estler, H.C.; Grewe, M.; Gausling, R.; Pavlovic, M.; Decker, K.
Biol. Chem. Hoppe-Seyler 373, 271-281, 1992
A/Title: Rat tumor necrosis factor-alpha. Transcription in rat Kupffer cells and in v
A/Reference number: S21674; M01D:92329007
A/Accession number: S21674
A/Molecule type: mRNA
A/Residues: 1-38, 'P', '40-162, 'T', '164-201, 'S', '203-235 <EST>
A/Cross-references: GB:X6539; GB:S40199; NID:g395369; PID:g395370
C/Comment: Tumor necrosis factor is secreted by macrophages in response to endotoxin
C/Genetics:
A/Gene: TNF-alpha
A/Intons: 62/3; 81/1; 97/1
C/Superfamily: tumor necrosis factor
C/Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
F;80-225/Product: tumor necrosis factor #status predicted <MAT>
F;86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 96; DB 2; Length 235;
Best Local Similarity 27.5%; Pred. No. 0.052;
Matches 58; Conservative 29; Mismatches 68; Indels 56; Gaps 15;

QY	22	GTGPGSONGEGYPMQSLPEQSSDALFEAMENGERSRKRRAVLTOKRKHNS---VLHLVP	77
DB	54	GVIGPKNE-EKFP-NGPLISSMA-----QTLTLRSSSSONSNDKRAHVVA	97
QY	78	INATSKSDVTEVMOPALRRGRLOAGYGVRA-----IQDAGVYLILYSVLFQD----	129
DB	98	NHQAE-----OLEW--LSORANALLANGDKLNDLVLPADGLYLYSVLEFKGGGCP	149
QY	129	--VFETMGVYSREGQGRQ--TLFRCLRS--MPSHDDRA-----YNSQVSAGVPPLHOG	177
DB	150	DYVLLT--HTVSRPAISYQERVSLSAISKPCPKDTEGAELEKWPYPMYGVFQLEKG	207
QY	178	DILS--VIIPRAKLNLSPHG--TFLEGVKL	205
DB	208	DLSAEVNLKP---YLDITESGVYFVIAL	235

RESULT 2

tumor necrosis factor - baboon
C:Species: Papio sp. (baboon)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text-change 08-Sep-1997
A:Accession: S22052
R:Sanjanwala, M.; Edwards, A.
Submitted to the EMBL Data Library, September 1991
A:Description: Baboon Tumor Necrosis Factor Derived from Sequences of Genomic DNA.
A:Reference number: S22052
A:Accession: S22052
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <S&N>
A:Cross-references: EMBL:X62141; NID:g38159; PID:g38160
C:Genetics:
A:Introns: 62/3; 78/1; 94/1
A:Superfamily: tumor necrosis factor
C/Keywords: transmembrane protein

Query Match 8.7%; Score 92; DB 2; Length 233;
Best Local Similarity 27.2%; Pred. No. 0.12;
Matches 41; Conservative % 23; Mismatches 51; Indels 36; Gaps 8;

OY 78 INANSKDSD--VTEVMNQPAL-----RRGRRLQAGYGVR-----IQDAGYLILYS 122
Db 77 VRSSRTPTSPDXPAHVAVNPQAEGOLQWLNRPANLLANGVELRDNLQVSEGLYLILYS 136
OY 123 QVLFPD-----VPTMGCV-VSRGCGROETLFR--CIRSWPSHPDBA--YNSQYS 168
Db 137 QVLFRGGCCPSTHYLTITISRIANSYOTRYNLSAISPCORETPDEGAEPKWTPEIYL 196
OY 169 AGVPHLHOGDLTSLVTIPRAKLNTSPHGTF 199
Db 197 GGVPQLKEKGDRLS-----AEINLPDYLDLF 220

RESULT 3
OWRN
tumor necrosis factor alpha precursor - human
A:Alternate names: cachectin; TNF alpha
C:Species: Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text-change 20-Mar-1998
A:Accession: A93585; S36153; A93351; A44189; B61478; I53311; S62610; A01646; B23784
R:Nedwin, G.E.; Naylor, S.L.; Sakaguchi, A.Y.; Smith, D.; Jarett-Nedwin, J.; Pennica, D.
Nucleic Acids Res. 13, 6361-6373, 1985
A>Title: Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromo-
some location
A:Reference number: A93585; MUID:86016093
A:Accession: A93585
A:Molecule type: DNA
A:Residues: 1-233 <NED>
A:Cross-references: GB:X02910; GB:X02159; NID:g37209; PID:g37210
R:Irits, F.J.M.; Bougueleret, L.; Priet, S.; Caterina, D.; Primas, G.; Perrot, V.; Jurkiewicz,
J.; Nature Genet. 3, 137-145, 1993
A>Title: Dense Alu clustering and a potential new member of the NF-kappaB family within a
human chromosome 1
A:Reference number: S36152
A:Accession: S36153
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-233 <TRI>
A:Cross-references: EMBL:Z15026; NID:g37211; PID:g37212
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1992
R:Pennica, D.; Nedwin, G.E.; Hayflick, J.S.; Seeburg, P.H.; Derynck, R.; Palladino, M.A.
Nature 312, 724-729, 1984
A>Title: Human tumour necrosis factor: precursor structure, expression and homology to lym-
phocyte activation factor
A:Reference number: A93351; MUID:8506244
A:Accession: A93351
A:Molecule type: mRNA
A:Residues: 1-233 <PEN>
A>Note: This protein was isolated from the monocyte-like cell line HL-60 from a promyeloid
leukemia patient
R:Many, A.M.; Creasey, A.A.; Ladner, M.B.; Lin, L.S.; Strickler, J.; Van Arsdale, J.N.;
Science 228, 149-154, 1995
A>Title: Molecular cloning of the complementary DNA for human tumor necrosis factor.

```

A:Reference number: A44189
A:Accession: A44189
A:Molecule type: mRNA
A:Residues: 1-62, 'S', '64-233 <MAN>
A:Cross-references: GB:M10988; NID:g339788
R:Fukuda, S.; Ando, S.; Sanou, O.; Tanial, M.; Fujii, M.; Masaki, N.; Nakamura, K.I.;
Lymphokine Res. 7, 175-185, 1998
A:Title: Simultaneous production of natural human tumor necrosis factor-alpha, -beta
A:Reference number: A61478
A:Accession: B61478
A:Molecule type: Protein
A:Residues: 83-102;109-119;121-128, 'X', '130-131;142-144, 'X', '146, 'xxx', '150-152;159-174;
Eur. J. Biochem. 152, 515-522, 1995
A:Title: Molecule cloning and expression of human tumor necrosis factor and comparis
A:Reference number: I53311; MUID:86030296
A:Accession: I53311
A:Status: translated from GB/EMBL/DDbJ
A:Molecule type: DNA
A:Residues: 1-233 <RES>
A:Cross-references: GB:M26331; NID:g339763; PID:g339764
A:Experimental source: U-937 cells
R:Takakura-Yamamoto, R.; Yamamoto, S.; Fukuda, S.; Kurimoto, M.
Eur. J. Biochem. 235, 431-437, 1996
A:Title: O-Glycosylated species of natural human tumor-necrosis factor-alpha.
A:Reference number: S62610
A:Accession: S62610
A:Status: preliminary
A:Molecule type: protein
A:Residues: 77-99 <TAK>
C:Comment: Secreted from mitogen-activated macrophages within 4-24 hours after induc
out detriment to normal cells. It can also act synergistically with interferon gamma
C:Comment: TNF-alpha and -beta (lymphotoxin) are the products of different genes clos
ut are produced by different cell types and have different induction kinetics.
C:Genetics:
A:Gene: GDB:TNF; TNFA
A:Cross-references: GDB:120441; OMIM:191160
A:Map position: 6p21.3-6p21.3
A:Introns: 62/3; 78/1; 94/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
F:77-233/Product: tumor necrosis factor #status experimental <MAT>
F:145-177/Disulfide bonds: #status experimental

Query Match      8.6%; Score 91.5; DB 1; Length 233;
Best Local Similarity 28.3%; Pred. No. 0.14;
Matches 39; Conservative 20; Mismatches 50; Indels 29; Gaps 7;

OY 78 INATSKDDSD--VTEVMQPAL-----RSGGLQAGYGVR----IQDAGVLLYS 122
      :.:.:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 VRSSSRTPSPDKPVAHVANVPAQESQLQWLNRRANALLANGVELRDNDLVVPSGLTLYIS 136
OY 123 QVLNQD-----VFETMGQV-VSREGGCRQETLR--CIRMPSPHPRD--YNSQYS 168
      |||||: | | | | | | | | | | | | | | | | | | | | | |
Db 137 QVLRGGGCGPSTHALLHTTISRINAVSYOTKVNLLSAIKSPQCRTEPGAEAKPYEPIYL 196
OY 169 AGVPHLHGGDILSYIPR 186
      || | | | | | | | | | | | | | | | | | | | | | |
Db 197 GGVFQLEKGRDLSEINR 214

RESULT 4
552715
tumor necrosis factor alpha precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 08-Sep-1997
;Accession: S52715

```


A:Title: Haploctypic polymorphisms of the TNFB gene.
A:Reference number: 154482; MUID:91139175
A:Accession: 154482
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-124, 'P', 126-205 <RES>
A:Cross-references: GB:M55913; NID:g339742; PID:g339743
A:Experimental source: ancestral haplotype 57.1
A:Note: 59-Asn was also found (ancestral haplotype 8.1)
R:Gray, P.W.; Aggarwal, B.B.; Benton, C.V.; Bringman, T.S.; Henzel, W.J.; Jarrett, J.A.; Nature 312, 721-724, 1984
A:Title: Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumour
A:Reference number: A93350; MUID:85086243
A:Accession: A93350
A:Molecule type: mRNA
A:Residues: 1-205 <GRA>
A:Cross-references: GB:X01393; NID:g34444; PID:g34445
A:Experimental source: lymphoblastoid cell line RPMI-1788
R:Goeddel, D.V.; Aggarwal, B.B.; Gray, P.W.; Leung, D.W.; Nedwin, G.E.; Palladino, M.A.; Cold Spring Harb. Symp. Quant. Biol. 51, 597-609, 1986
A:Title: Tumor necrosis factors: gene structure and biological activities.
A:Reference number: A32877; MUID:87217059
A:Accession: 832877
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 35-205 <GOE>
R:Kobayashi, Y.; Miyamoto, D.; Asada, M.; Obinata, M.; Osawa, T.
J. Biochem. 100, 727-733, 1986
A:Title: Cloning and expression of human lymphotoxin mRNA derived from a human T cell hy
A:Reference number: A91906; MUID:87057135
A:Accession: A91906
A:Molecule type: mRNA
A:Residues: 1-59, 'N', 61-205 <KOB>
A:Cross-references: GB:DD0102; NID:g219913; PID:d1000508; PID:g219914
A:Note: the authors translated the codon TAT for residue 156 as Thr and ACC for residue
R:Fukuda, S.; Ando, S.; Sanou, O.; Tanai, M.; Fujii, M.; Masaki, N.; Nakamura, K.I.; Ad
Lymphokine Res. 7, 175-185, 1988
A:Title: Simultaneous production of natural human tumor necrosis factor-alpha, -beta and
A:Reference number: A61478
A:Accession: A61478
A:Molecule type: protein
A:Residues: 56-79, 86-95, 'X', 97, 'X', 99, 119-151, 'XX', 154-162, 'X', 164, 'X', 166, 'X', 168, 'X', 1
R:Voigt, C.G.; Maurer-Fogy, I.; Adolf, G.R.
FEBS Lett. 314, 85-88, 1992
A:Title: Natural human tumor necrosis factor beta (lymphotoxin). Variable O-glycosylatidic
A:Reference number: S26951
A:Accession: S26951
A:Molecule type: protein
A:Residues: 35-59, 'N', 61-205 <VOI>
A:Note: 60-Thr was also found
R:Fukushima, K.; Watanabe, H.; Takeo, K.; Nomura, M.; Asahi, T.; Yamashita, K.
Arch. Biochem. Biophys. 304, 144-153, 1993
A:Title: N-linked sugar chain structure of recombinant human lymphotoxin produced by CHO
A:Reference number: S34742
A:Contents: annotation
C:Comment: secreted from mitogen-activated lymphocytes within 1-2 days after induction,
while having no detrimental effect on normal cells. It can also act synergistically with
C:Comment: This protein and TNF-alpha (tumor necrosis factor) are the products of differ
ical activities but are produced by different cell types and have different induction k
C:Genetics:
A:Gene: GDB:LTB; LT; TNFB
A:Cross-references: GDB:120442; OMIM:153440
A:Map position: 6p21.3-6p21.3
A:Introns: 33/3; 69/1
A:Note: the first intron occurs before the initiator codon
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; homotrimer; lymphokine; macrophage
E:1-34/Domains: signal sequence #status predicted <SIG>
E:35-205/Product: lymphotoxin #status predicted <MAP>
F:41/Binding site: carbohydrate (Thr) (covalent) (partial) #status experimental
F:96/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

Query Match      8.1%, Score 85.5; DB 1; Length 205;
Best Local Similarity 23.9%; Pred. No. 0.44;
Matches 42; Conservative 19; Mismatches 50; Indels 65; Gaps 9;

QY 68 KQHSVLHLVPIINATSK-----DDSDVTEVMQNALRG---RGLAQAGYGVRIQDAGV 117
      :||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 49 RQHPKMHILA--HSTLTKPAHLIGDPSKONSLLWRANTRDAFLQDFSSLNNSLLVPTSGI 106
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 118 YLLYEQVLFQDVTFTMGQVVSREGGQROETLFRICIRSMPP---SHPDRAYNSQYSAGVP-- 173
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 107 YFVYEQVYFSS-----GKAYSPPKA-----TSSPLYLAAHVYOLFQSSQYPRHPVLL 149
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 173 -----HLHQGDILSVI--IPRAKAKLNSPHTGLG 201
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 150 SSQKMYEPGLQEPMLHSMYHGAFAFLTQGDQLSTHTDSDIPH----LVYLSPTVEFG 201
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 10
QJ01344
tumor necrosis factor alpha precursor - horse
N:Alternate names: cachectin; TNF alpha
C:Species: Equus caballus (domestic horse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Sep-1997
C:Accession: QJ01344
R:Su, X.; Morris, D.D.; McGraw, R.A.
Gene 107, 319-321, 1991
A:Title: Cloning and characterization of gene TNF alpha encoding equine tumor necrosis
A:Reference number: QJ01344; MUID:92084125
A:Accession: QJ01344
A:Molecule type: DNA
A:Residues: 1-234 <SDU>
A:Cross-references: GB:M64087; NID:g164244; PID:g164245
C:Comment: This protein is an important proximal mediator of endotoxemia.
C:Genetics:
A:Gene: TNF-alpha
A:Introns: 62/3; 79/1; 95/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protease;
F:78-234/Product: tumor necrosis factor alpha #status predicted <TUD>

Query Match      8.1%, Score 85.5; DB 2; Length 234;
Best Local Similarity 27.3%; Pred. No. 0.52;
Matches 41; Conservative 14; Mismatches 38; Indels 57; Gaps 8;

QY 78 INATSKDSD--VTEVMQNALR-----RKGGLAQAGYGVRIQD-----AGYLLYS 122
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 78 LKSSRTSPSDKPAHVAVNPAQEGQLQMLSGRANALLANGVKLIDNQLVPLDGLYLIS 137
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 123 QVLFQDVTFTMGQVVSREGGQROET-----LFRICIRSMPSH---P 159
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 138 QVLF-----KGGCGPSTHVLTLTISRLAVSYSKVNLSSAKNS-PCHTESP 183
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 160 DRA-----YNSQYSAGVPHLHQGDILSVII 184
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 184 EQAEAKPWYEPYILGVPFQLEKGDQLSAEI 213
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

A:Title: Molecular cloning and expression in Escherichia coli of the cDNA coding for
A:Reference number: A25454; MUID:862191711
A:Accession: A25454
A:Molecule type: mRNA
A:Residues: 1-234 <ITD>
A:Cross-references: GB:M12845; NID:g165759; PID:g165760

tumor necrosis factor alpha precursor - rabbit
N:Alternate names: cachectin; TNF alpha
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Mar-1998
C:Accession: A25454; A25451; J50727
R:Itô, H.; Yamamoto, S.; Kuroda, S.; Sakamoto, H.; Kajihara, J.; Kiyota, T.; Hayashi,
DNA 5, 149-156, 1986

```

R:Rito, H.; Shirai, T.; Yamamoto, S.; Akira, M.; Kawahara, S.; Todd, C.W.; Wallace, R.B.
DNA 5, 157-165, 1986
A:Title: Molecular cloning of the gene encoding rabbit tumor necrosis factor.
A:Reference number: A25451; MUID:86219712
A:Accession: A25451
A:Molecule type: DNA
A:Residues: 1-234 <I12>
A:Note: this sequence differs from that shown in having a Gln inserted between residues
R:Shakhov, A.N.; Kuprash, D.V.; Azizov, M.M.; Jongeneel, C.V.; Nedospasov, S.A.
Gene 95, 215-221, 1990
A:Title: Structural analysis of the rabbit TNF locus, containing the genes encoding TNF
A:Reference number: JH0309; MUID:91065534
A:Accession: J50727
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-62, 'O', 63-234 <SHA>
A:Molecule type: DNA
A:Cross-references: GB:M60340; GB:M35326; NID:g165754; PID:g165756
C:Genetics:
A:Introns: 62/3; 80/1; 96/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein
F:1-81/Domain: signal sequence #status predicted <SIG>
F:82-234/Product: tumor necrosis factor #status predicted <MAT>

Query Match	8.1%;	Score 85.5;	DB 2;	Length 234;
Best Local Similarity	23.9%;	Pred. NO. 0.52;		
Matches 47;	Conservative 28;	Mismatches 53;	Indels 69;	Gaps 12;

```

QY      61 VLTQKOKKQSHVLL-----VYINTSK--DDSDVTEYMMQPAL-----RGRGL 103
Db      55 VIGQEQEESPNNLHLVNPVQAMQTLIRASRLSDKPLAHVYANPYEQGLQWLMSQRNAL 114
QY      104 QAQGVGVRIOD-----AGVYLLYSQVLEFDQVYFTMGQVVSREGQ----- 144
Db      115 LAN--GKMLDNGQLVPAVDGIXLYLSQVLF-----SGQCRRSYVLLTHTVSR 159
QY      144 -----ROETLFCIRSMPSH---PDR-----YNSQYSAQVPHLHOGDILSVITPRAR 189
Db      160 FAVSYPNKVNLLSAIKS--PCHRETPPEAEAPMAYEPYILGCVQLEKGDRLSTEVNQPE- 218
QY      190 KLNLSPHG--TFLGFKVL 205
Db      218 YLDLAESQGVYFGIIAL 234

```

RESULT 12
154490
tumor necrosis factor - white-footed mouse
C:Species: Peromyscus leucopus (white-footed mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Feb-1997
C:Accession: 154490
R:Crew, M.D.; Filipowsky, M.E.
Immunogenetics 35, 351-353, 1992
A:Title: Sequence of the tumor necrosis factor/cachectin (TNF) gene from Peromyscus leuc
A:Reference number: 154490; MUID:92218012
A:Accession: 154490
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-235 <RES>
A:Cross-references: GB:M59233; NID:g202506; PID:g202507
C:Genetics:
A:Gene: lTNF
A:Introns: 62/3; 81/1; 97/1
A:Superfamily: tumor necrosis factor

	Query Match	7.8%	Score 82.5;	DB 2;	Length 235;
	Best Local Similarity	27.9%	Pred. No. 1;		
	Matches	56;	Conservative	25;	Mismatches 65;
					Indels 55; Gaps 14;
OY	22 GTGGGQQNGEGPMSLDEQSSDALEAMENGRSKRRRAVLQOKKQHS---	VHLHP	77		
	: :	:	:		

```

Db      54 GVIGP-QREEFP--NNLPIIGSMA-----QTLTRSSQNSDDKPAHVH-  97

QY      78 INATSKDDSDVTEVMQPALRG-RGIAQAGYVR-----IODAGVLLYSQVLFQD---  129
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :
Db      97 --ANHQVDE---QLEW---LSRGANALLANGMDLKQDQVLPADGLYLVSQYLFPGQGC  148
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :

QY      129 -----VFVTMQ-VYSREGGGRQGTLR--CISMSPHPRR--YNSQVSAPPHLHOGD  178
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :
Db      149 SSYYLLTHTVSRFVNSYEDKNNLSAITSPCPKETPEGSELKPMYEPIYLGVGFQLEKGD  208
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :

QY      179 ILSVILIPARAKNLSPHGTF 199
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :
Db      209 RLS-----AEVNLPRKILDE 222
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 13
154491
cell surface antigen - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Feb-1997
C:Accession: 154491
R:Kirists, M.J.; Kunz, H.W.; Cortese Hassett, A.L.; Gill, T.J.
Immunogenetics 35 365-377, 1992
A:Title: Genomic DNA sequence and organization of a TL-like gene in the grc-c/c region
A:Reference number: 154491; MUID:92250119
A:Accession: 154491
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-392 <RES>
A:Cross-references: GB:M74822; NID:g205461; PID:g205462
C:Genetics:
A:Introns: 31/1, 121/1, 205/1, 297/1, 336/1, 347/1, 362/3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: surface antigen

	Query Match	7.6%	Score 81;	DB 2;	Length 392;	
	Best Local Similarity	22.0%;	Pred. NO. 2.8;			
	Matches	37;	Conservative	22;	Mismatches	61; Indels 48; Gaps 8
QY	34 PWOSLPEOSSDALLEAMENGERSKRRAVLTKOKKOHSLVLTLPINATSKDDSDTEVMW	93				
Dd	80 PW-----QEGBELFEWE-----QYHAKKTQDOQSARNLMMLVREYNKSMDFHT--LQW	127				
QY	94 QPALRRGGGLAQGYGVRIQAGVYLILYSQVLEFDV-----FTFMGQVVS	138				
Dd	128 Q-----QGCDVG-SDGRLLHWYDQLAFDGDVDHPTLNKLRFWTMTSTVAQISQ	175				
QY	139 REGQGRQETLFRCIRSMSPHPDRAINSQSYAGVPILH-----QGDI	179				
Dd	176 PELERLRND--NCSELLQKYPEREKERRLLRSPPAHYATHROPBPGGD	221				

RESULT 14
 D71262
 Probable hemolysin - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 07-Aug-1998
 C:Accession: D71262
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, F.
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete
 A:Reference number: A71250
 A:Accession: D71262
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-357 <COL>
 A:Cross-references: GB:AE001262; GB:AE000520; NID:g3323254; PID:g3323257
 A:Experimental source: strain Nicholas
 C:Genetics:
 A:Gene: TP0936

This Page Blank (uspio)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 1999, 14:43:39 ; Search time 25.8 Seconds

(without alignments)
213.266 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061
Sequence: 1 MVQLTQDTLOSLSRREVSRL.....RARAKNLNSPHGTFLGTVKL 205

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	9.0	235	1	TNFA_RAT
2	95.5	9.0	234	1	TNFA_CAVPO
3	92	8.7	233	1	TNFA_PAPSP
4	91.5	8.6	233	1	TNFA_HUMAN
5	90.5	8.5	235	1	TNFA_HUMAN
6	88	8.3	229	1	TNFA_RABIT
7	87.5	8.2	235	1	TNFA_CEREI
8	87	8.2	233	1	TNFA_MOUSE
9	86.5	8.2	233	1	TNFA_BOVIN
10	86	8.1	233	1	TNFA_FELCA
11	86	8.1	233	1	TNFA_MACMU
12	86	8.1	197	1	TNFA_RABIT
13	85.5	8.1	234	1	TNFA_HORSE
14	85.5	8.1	205	1	TNFA_HUMAN
15	83.5	7.9	233	1	TNFA_CANFA
16	82.5	7.8	235	1	TNFA_PERLE
17	80	7.5	234	1	TNFA_SHEEP
18	80	7.5	233	1	TNFA_TRIYU
19	78	7.4	281	1	FASL_HUMAN
20	78	7.4	232	1	TNFA_PIG
21	77.5	7.3	291	1	TRAI_MOUSE
22	77	7.3	504	1	BLK_HUMAN
23	77	7.3	193	1	TNFA_CAPHI
24	76	7.2	171	1	RS4_HALMA
25	75.5	7.1	484	1	FHL1_YEAST
26	74	7.0	496	1	IBMP_CERV
27	74	7.0	1063	1	PDR1_YEAST
28	73.5	6.9	638	1	FTSH_HELPF
29	73.5	6.9	202	1	TNFB_MOUSE
30	73	6.9	670	1	NTRC_SYNY3
31	72	6.8	726	1	EP2_METJA
32	72	6.8	360	1	RFL_ECOLI
33	72	6.8	1586	1	SN22_HUMAN
34	71.5	6.7	608	1	ATOS_ECOLI
35	71.5	6.7	1863	1	BRCL_HUMAN
36	71.5	6.7	337	1	OMB_NETLA
37	71.5	6.7	513	1	TTR_ECOLI
38	71	6.7	416	1	NGFR_CHICK
39	71	6.7	175	1	YCFE_ECOLI
40	71	6.7	378	1	ZP47_BRARE
41	70.5	6.6	325	1	IF34_HUMAN
42	70.5	6.6	204	1	TNFB_PIG
43	70.5	6.6	554	1	YJUK_ECOLI

ALIGNMENTS

44 70 6.6 361 1 COOH_RHORI
45 70 6.6 278 1 FASL_RAT

P31895 rhodospirill
P36940 rattus norv

RESULT 1

TNFA_RAT
ID TNFA_RAT STANDARD; PRT; 235 AA.
AC P16599;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-TESTIS;
RX MEDLINE: 94040766.
RA KWON J., CHUNG I.Y., BENVENISTE E.N.;
RL GENE 132:227-236(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA SHIRAI T., SHIMIZU N., HORIGUCHI S., ITO H.;
RL AGRI. BIOL. CHEM. 53:1733-1736(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92329007.
RA ESTLER H.C., GREWE M., GAUSSLING R., PAVLOVIC M., DECKER K.;
RL BIOL. CHEM. HOPPE-SEYLER 373:271-281(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA DECKER K.F.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE OF 1-231 FROM N.A.
RC TISSUE-TAIL;
RA KIRISITS M.J., VARDIMON D., KONZ H.W., GILL T.J. III;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC - FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC - SUBUNIT: HOMOTRIMER.
CC - SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC - PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC - DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC - SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X65399; G395370; -
DR EMBL: L00981; G205254; -
DR EMBL: D00475; G220921; -
DR EMBL: AJ002278; E1154446; -
DR EMBL: L19123; G310232; -
DR PIR: J00029; J00029.
DR PIR: S21674; S21674.
DR PIR: J00868; J00868.
DR HSSP: P01375; 1TNF.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 79
FT CHAIN 80 235 TUMOR NECROSIS FACTOR.


```

Query Match          8.7%: Score 92: DB 1: Length 233;
Best Local Similarity 27.2%: Pred. No. 0.1;
Matches 41: Conservative 23: Mismatches 51: Indels 36; Gaps 8;

OY 78 INATKDDSD--VFTEVMQPAL-----RRGRRLQAGCYVR-----IQDAGVYLLYS 122
Db 77 VRSSRTSPDKPVAHVANPAQEGOLQWLNRRANALLNAGVELRNQLVPSGGLYLLYS 136
OY 123 QVLFOD-----VFTEMGCV-VSRREGQGHQELFR--CIRSNPSPHDDRA--TNSQYS 168
Db 137 QVLFQGGQCPSTHVLTLTHTSRINASYQTKVNLLSAISPQGRLETPEGAEKAPWIEPYL 196
OY 169 AGVPHLHOGDILSVIIPRAKLNISPQGT 199
Db 197 GGVFOLKGRDRLS-----AEINLPDYLD 220

RESULT 4
ID TNEA_HUMAN STANDARD: PRT: 233 AA.
AC PO1375;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNEA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE: 87217060.
RA NEDOSPASOV S.A., SHAKHOV A.N., TURETSKAYA R.L., METT V.A.,
RA AZIZOV M.M., GEORGIEV G.P., KOROBKO V.G., DOBRYNIN V.N.,
RA FILIPPOV S.A., BYSTROV N.S., BOLDYREVA E.F., CHUVPILO S.A.,
RA CHIMAROV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RA COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE: 85086244.
RA PENNICA D., NEDWIN G.E., HAYFLICK J.S., SEEBURG P.H., DENYINCK R.,
RA PALADINO M.A., KOHR W.J., AGGARWAL B.B., GOEDEL D.V.;
RA NATURE 312:724-729(1984).
RN [3]
RP SEQUENCE FROM N.A.
RP MEDLINE: 85137898.
RA SHIRAI T., YAMAGUCHI H., ITO H., TODD C.W., WALLACE R.B.;
RA NATURE 313:803-806(1985).
RN [4]
RP SEQUENCE FROM N.A.
RP MEDLINE: 86016093.
RA NEDWIN G.E., NAVLOR S.L., SAKAGUCHI A.Y., SMITH D.H.,
RA JARRETT-NEDWIN J., PENNICA D., GOEDEL D.V., GRAY P.W.;
RA NUCLEIC ACIDS RES. 13:6361-6373(1985).
RN [5]
RP SEQUENCE FROM N.A.
RP MEDLINE: 85142190.
RA WANG A.M., CEESEY A.A., LADNER M.B., LIN L.S., STRICKLER J.,
RA VAN ARDELL J.N., YAMAMOTO R., MARK D.F.;
RA SCIENCE 228:149-154(1985).
RN [6]
RP SEQUENCE FROM N.A.
RP MEDLINE: 86030296.
RA MAMMENOT A., FRANSSEN L., TAVERNIER J., DER HEYDEN J., TIZARD R.,
RA KANASHIMA E., SHAW A., JOHNSON M.J., SEMON D., MUELLER R.,
RA RUSSCHAERT M.R., VAN VLIET A., FLERS W.;
RA EUR. J. BIOCHEM. 152:515-522(1985).
RN [7]
RP SEQUENCE FROM N.A.
RP MEDLINE: 93272029.
RA IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERINA D., PRIVAS G.,
RA PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA COHEN D.;

```

RL	NAT. GENET.	3:137-145(1993).
RN		
RP	X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).	
RX	MEDLINE; 89159409.	
RA	JONES E.Y., STUART D.I., WALKER N.P.;	
RL	NATURE 338:225-228(1989).	
RN		
RP	X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).	
RX	MEDLINE; 91193276.	
RA	JONES E.Y., STUART D.I., WALKER N.P.;	
RL	J. CELL SCI. SUPPL. 13:11-18(1990).	
RN		
RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).	
RX	MEDLINE; 90008932.	
RA	ECK M.J., SPRANG S.R.;	
RL	J. BIOL. CHEM. 264:17595-17605(1989).	
RN		
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MUTANT ARG-107.	
RX	MEDLINE; 91184128.	
RA	OSTADE X.V., TAVERNIER J., PRANGE T., FIEBS W.;	
RL	EMBO J. 10:827-836(1991).	
RN		
RP	MARISOXYLATION.	
RX	MEDLINE; 93018820.	
RA	STEVENSON F.T., BURSTEN S.L., LOCKSLEY R.M., LOVETT D.H.;	
RL	J. EXP. MED. 176:1053-1062(1992).	
CC	-1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA. IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS.	
CC	-1- SUBUNIT: HOMOTRIMER.	
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM.	
CC	-1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING.	
CC	-1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION.	
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.	
DR	EMBL; X02910; G37210; -.	
DR	EMBL; M16441; G339741; -.	
DR	EMBL; X01394; G37220; -.	
DR	EMBL; M10989; G339738; -.	
DR	EMBL; M26331; G339764; -.	
DR	EMBL; Z15026; G37212; -.	
DR	PIR; B23784; OMHUN.	
DR	PIR; A44189; A44189.	
DR	PIR; S36153; S36153.	
DR	PDB; 1TNF; 15-JAN-91.	
DR	PDB; 2TUN; 31-JAN-94.	
DR	PDB; 1ABW; 17-JUN-98.	
DR	MIM; 191160; -.	
DR	PROSITE; PS00251; TNF_1; 1.	
DR	PROSITE; PS50049; TNF_2; 1.	
KW	CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; MYRISTYLATION; 3D-STRUCTURE.	
KM		
FT	PROPEP	1
FT	CHAIN	77
FT	TRANSMEM	36
FT	LIPID	19
FT	LIPID	20
FT	DISULFID	145
FT	MUTAGEN	105
FT	MUTAGEN	108

FT MTAGEN 112 112 L->F: BIOLOGICALLY INACTIVE.
 FT MTAGEN 160 160 A->V: BIOLOGICALLY INACTIVE.
 FT MTAGEN 162 162 S->F: BIOLOGICALLY INACTIVE.
 FT MTAGEN 167 167 V->A,D: BIOLOGICALLY INACTIVE.
 FT MTAGEN 222 222 E->K: BIOLOGICALLY INACTIVE.
 FT CONFLICT 63 63 F -> S (IN REF. 5).
 FT TRN 85 86
 FT STRAND 89 94
 FT TRN 99 100
 FT STRAND 104 105
 FT TRN 109 110
 FT STRAND 112 113
 FT TRN 115 116
 FT STRAND 118 119
 FT STRAND 124 125
 FT STRAND 132 138
 FT STRAND 140 142
 FT STRAND 152 159
 FT STRAND 166 174
 FT STRAND 190 191
 FT STRAND 194 200
 FT STRAND 207 212
 FT HELIX 215 217
 FT STRAND 218 218
 FT TRN 224 225
 FT STRAND 226 232
 SQ SEQUENCE 233 AA; 25644 MW; 666D7069 CRC32;

Query Match 8.6%; Score 91.5; DB 1; Length 233;
 Best Local Similarity 28.3%; Pred. No. 0.11;
 Matches 39; Conservative 20; Mismatches 50; Indels 29; Gaps 7;

QY 78 INAKSDSD--VEVMQPAL-----RRKGLAQGVGVR-----IDAGYLLYS 122
 DB 77 VRSSRRPSPDKPAHVAVNPAEQOLOLNRANALLANGVELRDNLVPESEGILYLS 136
 QY 123 QVLRQD-----VFTMGQV-VSREGQROETLFR--CIRSMPSHPRDA--YNSQYS 168
 DB 137 QVLRKGGGCPSTHVLTLTIRIAVSOTKYNLSAISKPCOREPEGAEMKPTPEYTL 196
 QY 169 AGVPHLHQGDILSVIIPR 186
 DB 197 GGVPQLEKGRDLSEAINR 214

RESULT 5
 ID TNFA_RABIT STANDARD; PRT; 235 AA.
 AC P04924;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS ORCOTLAGUS CUNICULUS (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; LAGOMORPHA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91065534.
 RA SHAKHOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
 RA NEDOSPASOV S.A.; 215-221(1990).
 RL GENE 95:215-221(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86219712.
 RA ITO H., SHIRAI T., YAMAMOTO S., AKIRA M., KAWAHARA S., TODD C.W.,
 RA WALLACE R.B.; 5:157-165(1986).
 RL DNA 5:157-165(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86219711.

RA ITO H., YAMAMOTO S., KURODA S., SAKAMOTO H., KAJIHARA J., KIYOTA T.,
 RA HAYASHI H., KATO M., SEKO M.;
 RL DNA 5:149-156(1986).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CC CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.
 CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 CC AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL; M12845; G165760; -;
 DR EMBL; M12846; G165753; -;
 DR EMBL; M60340; G165756; -;
 DR PIR; A25451; A25451.
 DR PIR; A25454; A25454.
 DR PIR; J50727; J50727.
 DR HSSP; P01375; 1TNF.
 DR PROSITE; PS00251; TNF_1; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 79
 FT CHAIN 80 235 TUMOR NECROSIS FACTOR.
 FT TRANSEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 148 179 BY SIMILARITY.
 FT CONFLICT 63 63 MISSING (IN REF. 3).
 SQ SEQUENCE 235 AA; 25816 MW; 1BC5CDB CRC32;

Query Match 8.5%; Score 90.5; DB 1; Length 235;
 Best Local Similarity 24.1%; Pred. No. 0.14;
 Matches 48; Conservative 28; Mismatches 54; Indels 69; Gaps 12;

QY 59 RAVLTOKOKROHSHVHL-----VPINATSK--DSDVTEVMQPAL-----RRGR 101
 DB 54 RVIGPQEEQSPNNLHVNPAQMTVLRSAISRALSDKPLAHVAVNPAEQOLOLWSGRAN 113
 QY 102 GLQAGYGVRIQD-----AGVYLLYSQVLFQDVFTMGQVVSREG----- 144
 DB 114 ALLAN--GMKLTDLQVLVPADGLYLYSQVLF-----SQGCRSYVLLTHTV 158
 QY 144 -----ROETLFRCTIRSMPSH--PDRA-----YNSQYSAGVPHLHQGDILSVIIPRA 187
 DB 159 SRFAVSTPNKYNLSAISK--PCHRETPDEAPMAMVPEYTLGGVFOLEKGRDLSTEVNQP 217
 QY 188 RAKLNLSPHG-TFLGFYKL 205
 DB 218 E-YLDLAEISGVYRGITALL 235

RESULT 6
 ID TNFA_CEREL STANDARD; PRT; 229 AA.
 AC P51743;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN) (FRAGMENT).
 GN TNFA.
 OS CERVUS ELAPHUS (RED DEER).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; ARTIODACTYLA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LOCKHART E.A.;

```

RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; U14683; G540279; -.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT NON_TER 1
FT PROPEP 1
FT CHAIN 73 BY SIMILARITY.
FT TRANSEM 74 229 TUMOR NECROSIS FACTOR.
FT DISULFID 31 51 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 141 173 BY SIMILARITY.
SQ SEQUENCE 229 AA; 24987 MW; 4FD25D97 CRC32;

Query Match 8.3%; Score 88; DB 1; Length 229;
Best Local Similarity 28.8%; Pred. No. 0.24;
Matches 47; Conservative 24; Mismatches 40; Indels 52; Gaps 13;

QY 61 VLTKQKKQKHS---VLHLVP-INATSKDDSDVTEVMQPALRGRGLQAQGYRIOD- 115
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DQ 70 VQTLRSSQASINKPVAVVANINAG-----QLWLMDSC--ANALMAN--GVKLEDN 118
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 115 -----AGVYLKSYQLDQ-----VFTMGQV-VSNEGQROETLFRCLSPMSH 158
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DQ 119 QLVVETDLYLYSOVLRGQSCPSPTPLPLHTHTISRIASVY--TKVNIISAIRK-PC 174
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 159 ---PDRA-----YNSQYSGVPHLHQGDILSVILIRAKNL 193
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
DQ 175 RETPMAEAKPWEPYIYGVGFQLEKGRDL-----AEINL 210
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

RESULT 7
ID TNFA_MOUSE STANDARD: PRT: 235 AA.
AC P06804;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA OR TNF.
OS MUS MUSCULUS (MOUSE).
OC EUAROTIA; METACOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88224564.
RA SHIRAI T., SHIMIZU N., SHIOJIRI S., HORIGUCHI S., ITO H.;
RL DNA 7:193-201(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85298296.
RA PENNICA D., HAYFLICK J.S., BRINGMAN T.S., PALLADIO M.A.,
RA GOEDEL D.V.;
RL PROC. NATL. ACAD. SCI. U.S.A. 82:6060-6064(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 86149365.
RA CARUT D., BEUTLER B., HARTOG K., THAYER R., BROWN-SHIMER S.,

```

```

RA CERAMI A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:1670-1674(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85242112.
RA FRANSSEN L., MULLER R., MARMENOUT A., TAVERNIER J., VAN DER HEYDEN J.,
RA KAWASHIMA E., CHOLET A., TIZARD R., VAN HEUVERSWIN H., VAN VLIET A.,
RA RUTSCHAERT M.-R., FIEVS W.;
RL NUCLEIC ACIDS RES. 13:4417-4429(1985).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87298639.
RA SHAKHOV A.N., NEDOSPASOV S.A.;
RL BIOORG. KHIM. 13:701-705(1987).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88067722.
RA SEMON D., KAWASHIMA E., JONGENEEL C.V., SHAKHOV A.N., NEDOSPASOV S.A.;
RL NUCLEIC ACIDS RES. 15:9083-9084(1987).
RN [7]
RP SEQUENCE FROM N.A.
RX STRAIN-CTS, AND NOD;
RX MEDLINE: 96013654.
RA IREGAMI H., MAKINO S., YAMATO E., KAWAGUCHI Y., UEDA H., SAKAMOTO T.,
RA TAKEKAWA K., OGIHARA T.;
RL J. CLIN. INVEST. 96:1936-1942(1995).
RN [8]
RP SEQUENCE OF 80-99.
RX MEDLINE: 91097531.
RA SHERRY B., JUC D.-M., ZENTELLA A., CERAMI A.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 173:1072-1078(1990).
RN [9]
RP SEQUENCE OF 70-87.
RX MEDLINE: 89380231.
RA CSEH K., BEUTLER B.;
RL J. BIOL. CHEM. 264:16256-16260(1989).
RN [10]
RP IDENTIFICATION OF MEMBRANE-BOUND FORM.
RX MEDLINE: 88165056.
RA KRIEGLER M., PEREZ X., DEFRAY K., ALBERT I., LU S.D.;
RL CELL 53:45-53(1988).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; U06950; G495468; -.
DR EMBL; M13049; G302083; -.
DR EMBL; M11731; G302085; -.
DR EMBL; Y00467; G54832; -.
DR EMBL; X02611; G54845; -.
DR EMBL; M20155; G202093; ALT_SEQ.
DR EMBL; M38296; G202087; -.
DR EMBL; D84196; D1020271; -.
DR EMBL; D84194; D1020271; JOINED.
DR EMBL; D84195; D1020271; JOINED.
DR EMBL; D84197; D1020272; -.
DR EMBL; D84198; D1020272; JOINED.
DR PIR; A23127; OWMNS.
DR PIR; A22908; A22908.
DR PIR; A25164; A25164.

```

DR PIR: A27303; A27303.
 DR PIR: A34251; A34251.
 DR PIR: S03791; S03791.
 DR HSSP: P01375; 1TNF.
 DR MGD: MGT:104798; TNF.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT CHAIN 1 79
 FT PROPEP 1 79
 FT TRANSMEM 36 235
 FT DISULFID 148 179
 FT CARBOHYD 86 86
 FT CONFLICT 231 231
 FT SEQUENCE 235 AA; 25895 MW; 7FDE3C71 CRC32;

Query Match 8.2%; Score 87.5; DB 1; Length 235;
 Best Local Similarity 27.7%; Pred. No. 0.28;
 Matches 56; Conservative 22; Mismatches 67; Indels 57; Gaps 13;

QY 22 GTGCGPSQNEGYRWOSLPQSSDALAMWENGERSKRRRAVLTKOKKSHS---VLAHVP 77
 DB 54 GYIGP-QRDEKEP-NGCLPLISSMA-----QTLTLRSSQSSDKPVAHVYA 97
 QY 78 INATSKDSDVTEVMQPALRRGRGLQAGGYVR-----IODAGVYLXSOVLFOD---- 129
 DB 98 NHOVE-----QLEM--LSQRANALLANGMDLKNQVYPADGLVLYSOVLFKQGGCP 149
 QY 129 --VFTMGVSVREGGRDE--TLFRCLRS--MPSHPDRA---YNSQYSAQVPHLHOG 177
 DB 150 DYVLLT--HTVSFAISYQEKVNLTSVKSPCKDPPEGAELKPMYEPILYGVFQLENG 207
 QY 178 DILSVIIPRARKLNLSPHGT 199
 DB 208 DQLS-----AEVNLPKYLD 222

RESULT 8
 TNFA_BOVIN
 ID TNFA_BOVIN STANDARD; PRT; 233 AA.

AC 006599;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA OR TNF.
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ARTIODACTYLA.
 RN 11
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE; 94083525.
 RA CLUDIS I., CLEUTER Y., KETTMANN R., BURNY A., DROOGMANS L.;
 RL CYTOKINE 5:336-341(1993).
 RN 12
 RP SEQUENCE OF 50-233 FROM N.A.
 RC TISSUE-BLOOD;
 RX MEDLINE; 96006582.
 RA MERTENS B.E.L.C., MORIKURI M., GAUDULIS L.;
 RL IMMUNOGENETICS 42:430-431(1995).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.

CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 CC AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: Z14137; G798; -
 DR EMBL: Z48808; G755702; -
 DR PIR: S24642; S24642.
 DR HSSP: P01375; 1TNF.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT CHAIN 1 77
 FT PROPEP 1 77
 FT TRANSMEM 36 233
 FT DISULFID 145 177
 FT CONFLICT 62 62
 FT SEQUENCE 233 AA; 25439 MW; 207AFB2E CRC32;

Query Match 8.2%; Score 87; DB 1; Length 233;
 Best Local Similarity 25.5%; Pred. No. 0.3;
 Matches 51; Conservative 24; Mismatches 55; Indels 70; Gaps 13;

QY 19 RLQGTGSPQNEGYRWOSLPQSSDALAMWENGERSKRRRAVLTKOKKSHVLAHVP 78
 DB 60 REESGPGPSINS---PLVQTLRSSQA-----SNKPVAHVY-- 94
 QY 79 NATSKDSDVTEVMQPALRRGRGLQAGGYRID-----AGVYLXSOVLFOD--- 129
 DB 94 ---ADINSPGCLRWMDSY--ANALMAN--GVKLEDNOLVYPADGLVLYSOVLFKQGGC 145
 QY 129 ---VFTMGV--VSREGGRDETLFRCLRSMPH--PDRA---YNSQYSAQVPH 173
 DB 146 PSTPLFLHTLISRIAVSTQ---TKVNLISAKS-PCHRETPEMAEKPMYEPITGVGFQ 201
 QY 174 LHQGDILSVIIPRARKLN 193
 DB 202 LEKGDRLS-----AEINL 214

RESULT 9
 TNFA_FELCA
 ID TNFA_FELCA STANDARD; PRT; 233 AA.

AC P19101;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS FELIS SILVESTRIS CATUS (CAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; CARNIVORA.
 RN 11
 RN RP SEQUENCE FROM N.A.
 RX TISSUE-BLOOD;
 RX MEDLINE; 91016860.
 RA MCGRAW R.A., COFFEE B.W., OTTO C.M., DREWS R.T., RAWLINGS C.A.;
 RL NUCLEIC ACIDS RES. 18:5563-5563(1990).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE-BONE MARROW;
 RA DANIEL S.L., BRENNER C.A., LEGENDRE A.M., SOLOMAN A., ROUSE B.T.;
 RL SUBMITTED (XX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.

```
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X54000; G295777; -
DR EMBL: M82061; G403366; -
DR PIR: S11688; S11688.
DR HSSP: P01375; TNF.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
FT CONFLICT 28 28
FT CONFLICT 104 104
FT CONFLICT 151 151
FT CONFLICT 155 155
FT CONFLICT 210 210
SQ SEQUENCE 233 AA; 25322 MW; 1A5651A6 CRC32;

Query Match 8.2%; Score 86.5; DB 1; Length 233;
Best Local Similarity 24.4%; Pred. No. 0.34;
Matches 43; Conservative 22; Mismatches 76; Indels 35; Gaps 8;

QY 64 QKQKHSHVHLVPINAT-----SKDDSDVTEVMQPAL-----RRGRGLQAGY- 109
DB 59 QREELPHGLDLPLOTLLSSRTSPDKVAHVANPEAGOLQRLSRANALANGVE 118
QY 109 ----GVRIODAGRYLLYSQVLF-----QDVTFTMGQVVSREGSGRQETL-----FRCI 152
DB 119 LFDNQKVPDGLYLITYSQVLFQGQCPSTHVLTHAISFAVSQTKVNLSSAKSPCQ 178
QY 153 RSMPSHPDRA--YNSQVSAGVPHHOGDILSVIIPRARKLNSPHGTF 205
DB 179 RETPEGAENAPWEPIYLGCVFQLEKGRDLSTEL-NLPAYLDFAESGOVYFGIATL 233

RESULT 10
TNFA_MACFA STANDARD; PRT; 233 AA.
AC P79337;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOCYTES.
RA TATSUMI M.;
SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
```

```
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: AB000513; G1794149; -
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
SQ SEQUENCE 233 AA; 25558 MW; 856FC354 CRC32;

Query Match 8.1%; Score 86; DB 1; Length 233;
Best Local Similarity 26.5%; Pred. No. 0.38;
Matches 40; Conservative 23; Mismatches 52; Indels 36; Gaps 8;

QY 78 INATSKDDSD--VTEVMQPAL-----RRGRGLQAGYGR-----IQDAGVILYS 122
DB 77 VRSSRTSPDKVAHVANPEAGOLQRLSRANALVANGVELTDNQLVPSGSLTILYS 136
QY 123 QVLFQD-----VFTMGQV-VSREGGRQETLFR--CIRSMPSHPDRA--YNSQVS 168
DB 137 QVLFKQGCPSNHVLTHITSRIAVSQTKVNLSSAKSPCQETPEGAENAPWEPIYL 196
QY 169 AGVPHLHOGDILSVIIPRARKLNSPHGTF 199
DB 197 GGVFQLEKGRDLS-----AEINLPDYDF 220

RESULT 11
TNFA_MACMU STANDARD; PRT; 233 AA.
AC P48094;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS MACACA MULATTA (RHESUS MACAQUE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96003435.
RA VILLINGER F.J.; BRAR S.S.; MAYNE A.E.; CHIKKALA N.; ANSARI A.A.;
RL J. IMMUNOL. 155:3946-3954(1995).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: U19850; G644818; -
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
SQ SEQUENCE 233 AA; 25630 MW; 146A1B48 CRC32;
```

Query Match 8.1%; Score 86; DB 1; Length 233;
Best Local Similarity 26.5%; Pred. No. 0.38;
Matches 40; Conservative 23; Mismatches 52; Indels 36; Gaps 8;

QY 78 INATSKDSD--VTEVMQPAL-----RGRGLQAGYGV---TODAGVYLLYS 122
DB 77 VRSSRRPSPKPAHVANPQAEGLQWLNRRANALLANGVELDNDQNVPSGGLYIYS 136
QY 123 QVLEFD-----VFTMQGV-VSRGGGRQETLR--CIRSMPSHPRA--YNSOYS 168
DB 137 QVLEFGGCGPSNHVLLTHTTISRILAVSYQTKNLLSAIKSPCQRETPGAEAKPWTEPYTL 196
QY 169 AGVPHLHOGDILSYIIPRAKLNLSPHGTF 199
DB 197 GGVFLKRGDRLS-----AEINLPDYLD 220

RESULT 12
TNFA_RABBIT STANDARD; PRT; 197 AA.
AC P10154;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN LTA OR TNF.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91065534.
RA SHAKHOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
RA NEDOSPASOV S.A.;
RL GENE 95:215-221(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90220566.
RA SHAKHOV A.N., KUPRASH D.V., TURETSKAYA R.L., AZIZOV M.M.,
RA ANDREYEVA A.V., NEDOSPASOV S.A.;
RL MOL. BIOL. (MOSK) 23:1743-1750(1989).
CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC -1- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X55745; G297168; ALT_SEQ.
DR EMBL: M60340; G165755; -
DR EMBL: M60341; G165758; -
DR PIR: JH0309; JH0309.
DR PIR: P00098; P00098.
DR HSSP: P01375; 1TNF.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KM CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
FT SIGNAL 1 26
FT CHAIN 27 197 LYMPHOTOXIN-ALPHA.
FT CARBOHYD 88 88 POTENTIAL.
SQ SEQUENCE 197 AA; 21126 MW; 38A27DB CRC32;

Query Match 8.1%; Score 86; DB 1; Length 197;
Best Local Similarity 22.8%; Pred. No. 0.3;
Matches 49; Conservative 19; Mismatches 67; Indels 80; Gaps 10;

QY 26 PSQNGEGYPMOSLPEQSSDALEAMENGRSRRAVLTOKOKOHSYV---HLVPIINAT 81
DB 20 PPFGAQLPGAEFP-----SARNAQORLQAHFHSHTLKPAHLV----- 61
QY 82 SKDDSDVTEVMQPALRG---RGLQAGYGVRIODAGVYLLYSQVLFQDVTFTMGQVVS 138

DB 61 -GDPSQDSLNRMRANDRAFLRHGFLSNNSLLVSPSSGLYFYVSQVVE----- 108
QY 139 REGGRQETLFRCL-RKMP-----SHPDRAVNSQYSAGVPHL----- 176
DB 108 -SGEG-----CSKPAVPPLYLAHEVQLFSSQYSFHVBLLSNOKSVCPDGPQPVWVS 159
QY 176 -----OGDISVYIIPRAKLNLSPHGTF 201
DB 160 YQAGVFLITQGDQLSTHTD-GIAHLILSPSSVFFG 193

RESULT 13
TNFA_HORSE STANDARD; PRT; 234 AA.
AC P29553;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS EQUUS CABALLUS (HORSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; PERISSODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92084125.
RA SU X., MORRIS D.D., MCGRAW R.A.;
RL GENE 107:319-321(1991).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
CACHECTIN. IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
PROTEOLYTIC PROCESSING.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: M64087; G164245; -
DR PIR: J01344; J01344.
DR HSSP: P01375; 1TNF.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 77 BY SIMILARITY.
FT CHAIN 78 234 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 146 178 BY SIMILARITY.
SQ SEQUENCE 234 AA; 25469 MW; CD20BEE CRC32;

Query Match 8.1%; Score 85.5; DB 1; Length 234;
Best Local Similarity 27.3%; Pred. No. 0.42; Mismatches 38; Indels 57; Gaps 8;

QY 78 INATSKDSD--VTEVMQPALR-----RGRGLQAGYGVRIOD-----AGVYLLYS 122
DB 78 LRSSRRPSPKPAHVANPQAEGLQWLSGRANALLANGVELDNDQNVPLDGLYIYS 137
QY 123 QVLEFD-----VFTMQGV-VSRGGGRQET-----LFRCSMPSH---P 159
DB 138 QVLE-----KGGCGPSTHVLLTHTISRILAVSYKVNLLSAIKS-PCHTESP 183
QY 160 DRA-----YNSOYSAGVPHLHOGDILSVII 184
DB 184 EQAEAKPWTEPYIYIGVFLKRGDQLSAET 213

RESULT 14
TNFA_HUMAN

ID TNFB_HUMAN STANDARD: PRT: 205 AA.
AC P01374:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN LTA OR TNFB.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87217060.
RA NEDOSPASOV S.A., SHAKOV A.N., TURETSKAYA R.L., METT V.A.,
RA AZIZOV M.M., GEORGIEV G.P., KOROBKO V.G., DOBRININ V.N.,
RA FILIPPOV S.A., BYSTROV N.S., BOLDIREVA E.E., CHUPILLO S.A.,
RA CHUMAKOV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RA COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 86086150.
RA NEDIN G.E., JARRETT-NEDWIN J., SMITH D.H., NAYLOR S.L.,
RA SAKAGUCHI A.Y., GOEDEL D.V., GRAY P.W.;
RL J. CELL. BIOCHEM. 29:171-181(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87057135.
RA KOBAYASHI Y., MIYAMOTO D., ASADA M., OBINATA M., OSAMA T.;
RL J. BIOCHEM. 100:727-733(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85086243.
RA GRAY P.W., AGGARWAL B.B., BENTON C.V., BRINGMAN T.S., HENZEL W.J.,
RA JARRETT J.A., LEUNG D.W., MOFFAT B., NG P., SVEDERSKY L.P.,
RA PALLADINO M.A., NEDWIN G.E.;
RL NATURE 312:721-724(1984).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92339500.
RA MATSUYAMA N., OKAWA N., TSUKUI Y., ENDO T., KAWI A.;
RL FEBS LETT. 302:141-144(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93272029.
RA IRIS F.J.M., BOUGUELET L., PRIEUR S., CATERINA D., PRIMAS G.,
RA PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA COHEN D.;
RL NAT. GENET. 3:137-145(1993).
RN [7]
RP PARTIAL SEQUENCE, CARBOHYDRATE-BINDING SITES, AND VARIANT.
RX MEDLINE: 93083656.
RA VOIGT C.G., MAURER-FOGY I., ADOLF G.R.;
RL FEBS LETT. 314:85-88(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE: 92129275.
RA ECK M.J., ULTSCH M., RINDERNECHT E., DE VOS A.M., SPRANG S.R.;
RL J. BIOL. CHEM. 267:2119-2122(1992).
RN [9]
RP VARIANT ALLELE 8.1.
RX MEDLINE: 91139175.
RA ABRAHAM L.J., DU D.C., ZAHEDI K., DANKINS R.L., WHITEHEAD A.S.;
RL IMMUNOGENETICS 33:50-53(1991).
CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC -1- SUBUNIT: HETERODIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X01393; G34445; -
DR EMBL: X02911; G312411; -
DR EMBL: D00102; D1000508; -
DR EMBL: M16441; G339740; -

DR EMBL: D12614; G219912; -
DR EMBL: M55913; G339743; -
DR EMBL: Z15026; G37213; -
DR EMBL: A06316; G412161; -
DR PIR: A26085; OMHUX.
DR HSSP: P01375; 1TNB.
DR MIM: 153440; -
DR PROSITE: P500251; TNF_1; 1.
DR PROSITE: P500049; TNF_2; 1.
KW CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL; POLYMORPHISM.
FT SIGNAL 1 34
FT CHAIN 35 205 LYMPHOTOXIN-ALPHA.
FT CARBOHYD 41 41
FT CARBOHYD 96 96 PARTIAL.
FT VARIANT 60 60 T -> N.
FT VARIANT 125 125 T -> P (IN ALLELE 8.1).
FT CONFLICT 13 13 C -> R (IN REF. 6).
SQ SEQUENCE 205 AA; 22297 MW; 42F296D8 CRC32;

Query Match 8.1%; Score 85.5; DB 1; Length 205;
Best Local Similarity 23.9%; Pred. No. 0.36;
Matches 42; Conservative 19; Mismatches 50; Indels 65; Gaps 9;

QY 68 KQSVLHLVPINATSK-----DSDVTEVMQPALRG---RGLQAGCYVRDADAGV 117
DB 49 RQPKMHLA--HSTLKPAAHLIGDESKNSLIRANTDRAFLQDFSLSNSILVPTSGI 106
QY 118 YLLYSQVLEFDVTFMGQVNSREGGQROETLFCIRSNP---SHPDRAVNSQYAGVP-- 173
DB 107 YFYISQVYVS-----GKAISPKA-----TSSPLTYLAEHYQLFSSQYPPHVPILL 149
QY 173 -----HLHGGILSYI--IPRAAKLNLSPHGTFLG 201
DB 150 SSQKMYVPEIQEPWLSMYHGAFLQDQDLSHTHDGIPH---LVLSPTVTFEG 201

RESULT 15
ID TNFA_CANFA STANDARD: PRT: 233 AA.
AC P51742; 028339;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA OR TNF.
OS CANIS FAMILIARIS (DOG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUHERIA; CARNIVORA.
RN [1]
RP SEQUENCE FROM N.A.
RA FIERS W., BEERNERT M.;
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA ZUCKER K., LU P., FULLER L., ASTHANA D., ESCUENAZI V., MILLER J.;
RL LYMPHOKINE RES. 13:191-196(1994).
RN [3]
RP SEQUENCE OF 74-205 FROM N.A.
RC STRAIN-BEAGLE; TISSUE-BLOOD.
RA GILMORE W.H., CARTER S.D., BENNETT M., BARNES A., KELLY D.F.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CACHECTIN, IT IS A POTENT PROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY

```
CC      PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC      -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC      CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC      AND MALNUTRITION.
CC      -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR      EMBL: X94932; E222187; -.
DR      EMBL: S74068; G802045; -.
DR      EMBL: Z70046; E228213; -.
DR      PROSITE: PS00251; TNF_1; 1.
DR      PROSITE: PS50049; TNF_2; 1.
KM      CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT      PROPER 1
FT      CHAIN 77
FT      TRANSMEM 36
FT      DISULFID 145
FT      CONFLICT 59
FT      CONFLICT 66
FT      CONFLICT 74
FT      CONFLICT 111
FT      CONFLICT 116
FT      CONFLICT 134
SO      SEQUENCE 233 AA; 25447 MW; BCA64FA0 CRC32;

Query Match 7.9%; Score 83.5; DB 1; Length 233;
Best Local Similarity 24.7%; Pred. No. 0.65;
Matches 43; Conservative 26; Mismatches 62; Indels 43; Gaps 9;

QY 64 QKOKKSHVHLIVINATSK-----DDSDYTEVMQPAL-----RRGRGLQAQGYG 109
DB 59 QRELPNGLOLISPLAQTVKSSRTPSDKPYAHVYANPEAGQLQWLRRANALLANGVE 118
QY 110 VR-----IODAGVYLXYSQVLFOD-----VFTMGQ-VSREGGQROETLFR--CI 152
DB 119 LITDNLIVPSGGLITISQVLFKGGCGPSTHVLTHITISRAVSYQTKVNLISAIKSPCQ 178
QY 153 RSMSPHPDRA--YNSQYSAGVPHLHQGDILSVIIPRAKILNLSPHGTFIGFVR 204
DB 179 RETPEGTETAKPWEPYILGIVGFQLEKGDRLS-----AEINLP--NYLDFAE 222
```

Search completed: May 19, 1999, 14:43:40
Job time: 47 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:58 ; Search time 48.52 Seconds

(without alignments)

233.093 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061

Sequence: 1 MVOLTOQTELOSRLREVSR.....RAKALNLSFPHGLFVKKL 205

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPTREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_phase:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	97.2	250	2 075888	075888 homo sapien
2	107.5	10.1	391	2 075910	075910 homo sapien
3	90.5	8.5	157	2 043647	043647 homo sapien
4	89.5	8.4	377	10 054693	054693 mus musculi
5	87.5	8.2	583	8 043143	043143 lycopersico
6	87	8.2	233	4 018779	018779 bos taurus
7	87	8.2	1726	10 062383	062383 mus musculi
8	86	8.1	07510	2 07510	07510 papio hamad
9	85.5	8.1	1375	3 094537	094537 dirosophila
10	85.5	8.1	1526	3 094538	094538 dirosophila
11	85.5	8.1	364	10 054910	054910 mus musculi
12	85.5	8.1	319	12 093529	093529 xenopus lae
13	83	7.8	754	11 089250	089250 soil-borne
14	82	7.7	232	10 035853	035853 mus musculi
15	81	7.6	392	5 031277	031277 rattus norv
16	81	7.6	337	10 035734	035734 marimota mon
17	80.5	7.6	377	5 031279	031279 rattus norv
18	80.5	7.6	357	9 083906	083906 treponema p
19	80	7.5	1603	2 015737	015737 homo sapien
20	80	7.5	2957	3 061845	061845 caenorhabd1
21	80	7.5	513	9 054427	054427 salmonella
22	80	7.5	707	9 005415	005415 clostridium
23	79.5	7.5	234	4 028320	028320 capra hircu
24	79	7.4	319	9 048367	048367 enterococu
25	79	7.4	307	9 087203	087203 lactococcu
26	79	7.4	2629	10 P97499	P97499 mus musculi
27	79	7.4	1888	10 088466	088466 mus musculi
28	78.5	7.4	216	10 070332	070332 mesocricetu
29	78.5	7.4	842	10 008874	008874 rattus norv

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	250 AA.
075888				
AC	075888			
DT	01-NOV-1998 (TREMBLREL. 08, CREATED)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)			
DE	PROLIFERATION INDUCING LIGAND APRIL.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-UTERUS.			
RX	MEDLINE; 98416181.			
RA	HAHNE M., KATACKA T., SCHROETER M., HOFMANN K., IRMLER M.,			
RA	BODMER J.-L., SCHNEIDER P., BORNAND T., HOLLER N., FRENCH L.E.,			
RA	SORDAT B., RIMOLDI D., TSCHOEP J.			
RT	"APRIL, a New Ligand of the Tumor Necrosis Factor Family, Stimulates			
RT	Tumor Cell Growth."			
RL	J. EXP. MED. 188:1185-1190(1998).			
DR	EMBL; AF046888; G3650492; -.			
SQ	SEQUENCE 250 AA; 27433 MW; 48C59C35 CRC32;			
Query Match	97.28; Score 1031; DB 2; Length 250;			
Best Local Similarity	98.08; Pred. No. 3.7e-85;			
Matches 201; Conservative	0; Mismatches 4; Indels 0; Gaps 0;			
QY	1 MVOLTOQTELOSRLREVSRLOGTGPSONEGYPMOSLPPOSSDALPAMWENGSRKRA 60			073628 anolis caro
DB	46 MALLTOQTELOSRLREVSRLOGTGPSONEGYPMOSLPPOSSDALPAMWENGSRKRA 105			016291 homo sapien
QY	61 VLAGOKRKHSHVLAHPINATSKDSDVTEVMQPALRRGGLAOAGYVRIODAGYLL 120			022516 caenorhabd1
DB	106 VLFQKOKKSHVLAHPINATSKDSDVTEVMQPALRRGGLAOAGYVRIODAGYLL 165			018794 caenorhabd1
QY	121 YSQLFQDVTFTKQVYSREGQROETLFCIRKMPSPHPRAYSQYSAGVPHLHQDIL 180			047432 escherichia
DB	166 YSQLFQDVTFTKQVYSREGQROETLFCIRKMPSPHPRAYNSCYAGVFHLHQDIL 225			062326 mus musculi
QY	181 SVTPRARKLNLSPHGTFLGFVKL 205			090765 gallus gall
DB	226 SVTPRARKLNLSPHGTFLGFVKL 250			091493 caenorhabd1
RESULT	2			
075910				
ID	075910	PRELIMINARY:	PRT:	391 AA.
AC	075910			
DT	01-NOV-1998 (TREMBLREL. 08, CREATED)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)			

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ECTODYSPLASIN-A EDA ISOFORM II.
GN ED1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 98349961.
RA MONREAL A.W., ZONANA J., FERGUSON B.;
RT "Identification of a new splice form of the ED1 gene permits
RT detection of nearly all x-linked hypohidrotic ectodermal dysplasia
RT mutations.";
RL AM. J. HUM. GENET. 63:380-389(1998).
DR EMBL: AF060999; G3639041; -;
SQ SEQUENCE 391 AA; 41294 MW; 3DD4B2F9 CRC32;

Query Match 10.1%; Score 107.5; DB 2; Length 391;
Best Local Similarity 24.5%; Pred. No. 0.039;
Matches 47; Conservative 31; Mismatches 79; Indels 35; Gaps 7;

OY 22 GTGCGSNGEGTIPWOSLPEQSSDALEAMENGERSRRRAVITQOKOKHSHLVLPINAT 81
DB 221 GPPGG-----OGPPGLQSGSGADKAGTREN-----OPAVVHLQGGCSA 259
OY 82 --SKD-----SDTYEVMOPALRRGRGLQAGYGVRIQDAGVYLLYSQVLFQDVTFTM 133
DB 260 IOVKNDLSGGVINDMSRTIMNKVKR--LHPRSGELEVDGTFTYSQVEVYINFTD 316
OY 134 GGVNSREGGROETLFCIRSMPSHPDRAYNSQYAGVPHLQGDILSYIIPRAKAKLNL 193
DB 316 -DEAYEVVVDKPPQLQCTRSIETGKTN-YNTCYTAGVCLLAKRQIAVKVHADISIMM 373
OY 194 SPHGTFLEFVKL 205
DB 374 SKHTTFFGAIRL 385

RESULT 3
OY 043647 PRELIMINARY; PRT: 157 AA.
AC 043647;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA (FRAGMENT).
GN TNFA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA JANG J.S., KIM B.E.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF043342; G2905634; -;
DR PROSITE: PS00251; TNF.1; 1.
FT NON_TER
SQ SEQUENCE 157 AA; 17380 MW; FD78F0A6 CRC32;

Query Match 8.5%; Score 90.5; DB 2; Length 157;
Best Local Similarity 28.3%; Pred. No. 0.42;
Matches 39; Conservative 20; Mismatches 50; Indels 29; Gaps 7;

OY 78 INATSKDSD--VTEVMOPAL-----RRGRGLQAGYGV-----IQDAGVYLLYS 122
DB 1 VSSSRTPSDKVAHVAVNAPQAEGLQWLMRRANALLANGVELRDQNLVYPSGGLYLIYS 60
OY 123 QVLFQD-----VTFMGV-VSRREGQGOETLFR--CIRSMPSHPDRA--YNSQYS 168
DB 61 QVLFQGGCGPSTHVLTLTIRISRIAVSYQTKVNLISAIKSPCORETPRGAARPWYEPYL 120

OY 169 AGVPHLQGDILSYIIPR 186
DB 121 GGVPLEKGDRLSALINR 138

RESULT 4
OY 054693 PRELIMINARY; PRT: 377 AA.
AC 054693;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE EDA PROTEIN HOMOLOG.
GN TABBY.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIROGNATHI; MORIDAE; MORINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 97449184.
RA FERGUSON B.M., BROCKDORFF N., FORMSTONE E., NGUYEN T., KRONMILLER J.E.,
RA ZONANA J.;
RT "Cloning of Tabby, the murine homolog of the human EDA gene: evidence
RT for a membrane-associated protein with a short collagenous domain.";
RL HUM. MOL. GENET. 6:1589-1594(1997).
RN [2]
RP SEQUENCE OF 1-132 FROM N.A.
RC STRAIN-129;
RA BROCKDORFF N., ZONANA J., FORMSTONE E., FERGUSON B.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF004435; G2660700; -;
DR EMBL: AF004434; G2660698; -;
SQ SEQUENCE 377 AA; 40002 MW; AEA7ED8 CRC32;

Query Match 8.4%; Score 89.5; DB 10; Length 377;
Best Local Similarity 25.6%; Pred. No. 1.5;
Matches 33; Conservative 26; Mismatches 49; Indels 21; Gaps 5;

OY 94 OPALRRGRGLQAGYGVRIQD---AGVYLLYSQVLFQDVTFTM---GOV----- 137
DB 247 OPVVH--LOGGSGAIGVKNDLSGGVINDMSRTIMNKVKFRLHPRSGELEYYINFTDF 303
OY 137 VSRREGGROETLFCIRSMPSHPDRAYNSQYAGVPHLQGDILSYIIPRAKAKLNLSPH 196
DB 304 ASYEYVVDKPPQLQCTRSIETGKTN-YNTCYTAGVCLLAKRQIAVKVHADISIMSKH 362
OY 197 SPHGTFLEFVKL 205
DB 363 TTFGAIRL 371

RESULT 5
OY 043143 PRELIMINARY; PRT: 583 AA.
AC 043143;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PECTINSTERASE (EC 3.1.1.11) (PECTIN METHYLESTERASE) (PE).
GN PMEUL.
OS LYCOPERSTON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPLANTAE; CHAROPHYTA/EMBRYOPHYTA GROUP; EMBRYOPHYTA;
OC TRACHEOPHYTES; EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA;
OC EUDICOTYLEDONS; ASTERIDAE; SOLANANAE; SOLANALES; SOLANACEAE; SOLANUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO GRANDE;
RA GARFE J., TIZNADO M.E., HANDA A.K.;
RL PLANT PHYSIOL. 110:1436-1436(1996).
DR EMBL: U49330; G1222552; -;
DR PROSITE: PS00503; PECTINSTERASE_2; 1.

DB 137 QVLFKGGCGCPNHLTHTTSRIVSYOTKVNLSAISKPCQRTPEGAACKPWEPYIYL 196
 QY 169 AGVPHLHQGDILSYIPRAKAKLNSPHGTF 199
 DB 197 GGVFOLEKGRDLS-----AEINLPDYLD 220

RESULT 9
 094537
 ID 094537 PRELIMINARY; PRT; 1375 AA.

AC 094537;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE FRAZZLED.
 GN FRAZZLED.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 97015076.
 RA KOLODZIEJ P.A., TIME L., MITCHELL K.J., GOODMAN C.S., FRIED S.,
 RA JAN L.Y., JAN Y.N.;
 RT "frazzled encodes a Drosophila member of the DCC immunoglobulin
 RT subfamily and is required for CNS and motor axon guidance."
 RL CELL 87:197-204(1996)
 DR EMBL: U71001; G1621117; -
 DR FLYBASE; FBgn0011592; fra.
 DR PFAM; PF00041; fn3; 6.
 DR PFAM; PF00047; 19; 3.
 SO SEQUENCE 1375 AA; 151692 MW; 43806DBC CRC32;

Query Match 8.1%; Score 85.5; DB 3; Length 1375;
 Best Local Similarity 22.5%; Pred. No. 17;
 Matches 41; Conservative 13; Mismatches 47; Indels 81; Gaps 7;

QY 6 QQTELOSL-----REVSRLOGTG-----GPSQNGEGYPMOSLPE 40
 DB 516 QQVNIOSLPGRTYQFVEANTNFGSGASSAPLEVSTQPEVNINAGPPRNEG Y-----AR 570
 QY 41 QSSDALEAME-----NGERSRKR-----RAVLQKKOKKSHVILHY 76
 DB 571 SHKEIYKWEPTVTNGEILIKRYVYSENDGADLYHDSTALEAVLTLPHTDYISV 630
 QY 77 PINATSKDSD-----DYTE-----VMMOPALRRGRGLQAQGYG 109
 DB 631 PFNRNGMGDSAEIRVKTFSTPSEPNNVTLEVITSSSITVHWPPEAEDRNGQITGYK 690
 QY 110 VR 111
 DB 691 IR 692

RESULT 10
 094538
 ID 094538 PRELIMINARY; PRT; 1526 AA.

AC 094538;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE FRAZZLED.
 GN FRAZZLED.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 97015076.

RA KOLODZIEJ P.A., TIME L., MITCHELL K.J., GOODMAN C.S., FRIED S.,
 RA JAN L.Y., JAN Y.N.;
 RT "frazzled encodes a Drosophila member of the DCC immunoglobulin
 RT subfamily and is required for CNS and motor axon guidance."
 RL CELL 87:197-204(1996)
 DR EMBL: U71002; G1621117; -
 DR FLYBASE; FBgn0011592; fra.
 DR PFAM; PF00041; fn3; 6.
 DR PFAM; PF00047; 19; 3.
 SO SEQUENCE 1526 AA; 168787 MW; F17B1EC9 CRC32;

Query Match 8.1%; Score 85.5; DB 3; Length 1526;
 Best Local Similarity 22.5%; Pred. No. 20;
 Matches 41; Conservative 13; Mismatches 47; Indels 81; Gaps 7;

QY 6 QQTELOSL-----REVSRLOGTG-----GPSQNGEGYPMOSLPE 40
 DB 667 QQVNIOSLPGRTYQFVEANTNFGSGASSAPLEVSTQPEVNINAGPPRNEG Y-----AR 721
 QY 41 QSSDALEAME-----NGERSRKR-----RAVLQKKOKKSHVILHY 76
 DB 722 SHKEIYKWEPTVTNGEILIKRYVYSENDGADLYHDSTALEAVLTLPHTDYISV 781
 QY 77 PINATSKDSD-----DYTE-----VMMOPALRRGRGLQAQGYG 109
 DB 782 PFNRNGMGDSAEIRVKTFSTPSEPNNVTLEVITSSSITVHWPPEAEDRNGQITGYK 841
 QY 110 VR 111
 DB 842 IR 843

RESULT 11
 054910
 ID 054910 PRELIMINARY; PRT; 364 AA.
 AC 054910;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE NUCLEAR FACTOR OF KAPPA LIGHT POLYPEPTIDE PROTEIN ENHANCER IN B-CELLS
 DE INHIBITOR, EPSILON (IKB EPSILON).
 GN NFKBIE.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCIUOGNATHI; MORIDAE; MORINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 98070758.
 RA TIMEONDIS S., LIANG S., CHEN G., THANOS D.;
 RT "Cloning and functional characterization of mouse IkappaBepsilon."
 RL PROC. NATL. ACAD. SCI. U.S.A. 94:14372-14377(1997).
 DR EMBL; AF030896; G2739158; -
 DR MGI; MGI:1194908; NFKBIE.
 SO SEQUENCE 364 AA; 39327 MW; EC4923B7 CRC32;

Query Match 8.1%; Score 85.5; DB 10; Length 364;
 Best Local Similarity 23.1%; Pred. No. 3; 3;
 Matches 46; Conservative 30; Mismatches 82; Indels 41; Gaps 9;

QY 10 LOSLR--EVSRLQGTGPSQNGEGYPMOSLPPQSSDALEAMENGERSKRRRAVLQKK 67
 DB 23 LRSLSRSLPEPTAAGSGS--SQSGCPQFWRHAPETHEPEKEDADGERADSTYA----- 75
 QY 68 KQSVLHLVI--NATSKDSDVTEVMMOPALRRG-----RGLQAQGYVRIODAGVILY 121
 DB 75 -SSSLTSEFPLERPEAKDS--PVPVGSPLPAGVLSPOQLALTY--TSEGDGTLH 128
 QY 122 SQVLFQDVTFTMGQVVSREGQROETLFRICIRSPSHPRAYNSQYSAGV--PHLHQD 178
 DB 129 LAVIHE-----APSVLFCCLAFLPQEVLDIQNNKLTQTAHLAVHLDQDP 172

OY 179 ILSVIIPARAKLNSPHG 197
 Db 173 VVRAVLKGSASRILODQH 191

RESULT 12

093529 PRELIMINARY; PRT; 319 AA.
 AC 093529;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE FORKHEAD-DOMAIN-CONTAINING PROTEIN 5.
 GN FKH-5.
 OS XENOPUS LAEVIS (AFRICAN CLAMMED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 CC MESOBATRACHIA; PIPOIDEA; PIPOIDAE; XENOPODINAE; XENOPUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GANSE J.T., KUO J., PATEL M., SIVE H.;
 RT "Patterning of the xenopus ectoderm during blastula and gastrula stages."
 RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF064810; G3695057; -;
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 SQ SEQUENCE 319 AA; 34954 MW; B2A80638 CRC32;

Query Match 8.1%; Score 85.5; DB 12; Length 319;
 Best Local Similarity 23.4%; Pred. No. 2.8;

Matches 43; Conservative 23; Mismatches 85; Indels 33; Gaps 8;

OY 26 PSQNGEGYPMOSLPEQSSDALEAWNGERSKRRRAVLTKQKKQSHVHLVPINATSKD 85
 Db 78 PDQPKGSEFALHPRCG----DMFENGSEFLRRKRFKWKSD-----HLAPSKA----- 123
 OY 86 SDVTEVMQPALRRGRGLQAGQGYRIODAGYLL-YISQVLEFQDYFTMGQVVSREGGR 144
 Db 123 SDAQVILQQAQKLRLSALAAS--GTHLPMSMYNLGVSPTSSEFKRPFALIENTIAREYKMP 180
 OY 145 QETLEFRCIRSMPSHPRAY-----NSQYSAGVPHLHOGDILSVIIP--RAKAKLNL 193
 Db 181 GGLAESTMQMPMA---AYPLHMLTITVGSGITGMPHMTSSMLDSTTPIISMANSYSV 236
 OY 194 SPHG 197
 Db 237 SAYG 240

RESULT 13

089250 PRELIMINARY; PRT; 754 AA.
 AC 089250;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE 84 KDA READTHROUGH PROTEIN.
 OS SOIL-BORNE WHEAT MOSAIC VIRUS.
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FUDROVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-US-NEBRASKA, 1988 WILD-TYPE;
 RX MEDLINE: 93303914.
 RA SHIRAKO Y., WILSON M.A.;
 RT genome of soil-borne wheat mosaic virus."
 RL VIROLOGY 195:16-32(1993).
 DR EMBL: L07938; G295512; -;
 SQ SEQUENCE 754 AA; 83736 MW; B41CB937 CRC32;

Query Match 7.8%; Score 83; DB 11; Length 754;
 Best Local Similarity 26.1%; Pred. No. 14;
 Matches 46; Conservative 24; Mismatches 68; Indels 38; Gaps 9;

OY 13 LRREVSRLQGTGSPQNGEGYPMQSLPEQS-----SDALEAWNG-ERSKRRAY 61
 Db 580 LLRALSMLEPSYGGGGGNGGG---GSLPPQALELFPRACTFEERLALQNGDLSPEMNEV 636
 OY 62 LTQKKQSHVHLVPINATSKDSDVTEVMQPALRRGRGLQAGYVRI----QDAG 116
 Db 637 FTPEELKT-EIRHVYQAAVDSS---YHVADEPYLR-----GVGVTYVGTGLPEAVG 682
 OY 117 -VYLLSYVLQDYFTMGQVVSREGGRQETLFCISMSHPRAANSQYSAGV 171
 Db 683 EITTTGSESVSDVTSISLGV---PGRARRRAVSXNTHTMPTSGQSQSYNSKSPRY 734

RESULT 14

035853 PRELIMINARY; PRT; 232 AA.
 AC 035853;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR ALPHA.
 GN TNFA.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 CC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-A/J;
 RX MEDLINE: 97246744.
 RA IRAQI F., TEALE A.;
 RT "Cloning and sequencing of the tnfa genes of three inbred mouse strains."
 RL IMMUNOGENETICS 45:459-461(1997).
 DR EMBL: U68414; G2304957; -;
 DR PROSITE; PS00251; TNF_1; 1.
 DR PFAM: PF00229; TNF; 1.
 SQ SEQUENCE 232 AA; 25513 MW; 9B2B3F06 CRC32;

Query Match 7.7%; Score 82; DB 10; Length 232;
 Best Local Similarity 26.3%; Pred. No. 3.9;
 Matches 52; Conservative 24; Mismatches 70; Indels 52; Gaps 12;

OY 22 GTGSPQNGEGYPMOSLPEQSSDALEAWNGERSKRRRAVLTKQKKQSHVHLVPINAT 81
 Db 54 GVTGP-QRDEKFP-NGLEPLISSMA-----QTLSSQNSSDKPAHAVVANHQV 98
 OY 82 SKDSDVTEVMQPALRRGRGLQAGYVRI----QDAGYLLTQVLEFOD-----VT 130
 Db 99 EE-----QLEM--LSQANALLANGMDLKNQVLPADGLITVYSLVLFKGGCGPDYVL 150
 OY 131 FTMGQVVSREGGRQETLFCIRSR--MPSHPDRA-----YNSQYSAGVPHLHOGDILS 181
 Db 151 LT--HTVSREFAISYQEKVNLISAVKSPCKPTPEGAELKPWEPIYLLGVPQLERGDLS 208
 OY 182 VIIPARAKLNSPHGTF 199
 Db 209 -----AEVNLPKYIDF 219

RESULT 15

031277 PRELIMINARY; PRT; 392 AA.
 AC 031277;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE CELL SURFACE ANTIGEN PRECURSOR.
 OS RATTUS NORVEGICUS (RAT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92250119.
RA KIRISITS M.J., KUNZ H.W., CORTESI HASSETT A.L., GILL T.J. III;
RT "Genomic DNA sequence and organization of a TL-like gene in the
RL IMMUNOGENETICS 35:365-377(1992).
DR EMBL: M74822; G205462; -
DR PFAM: PF00047; 1g; 1.
DR PFAM: PF00129; MHC_I; 1.
KW SIGNAL; MHC.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 392 CELL SURFACE ANTIGEN.
SQ SEQUENCE 392 AA; 44744 MW; 9C2990D8 CRC32;

Query Match 7.6%; Score 81; DB 5; Length 392;
Best Local Similarity 22.0%; Pred. No. 9.2;

Matches 37; Conservative 22; Mismatches 61; Indels 48; Gaps 8;

QY 34 PMQSLEQSSDALFEWENGRSRKRAVLTQKQKQSHVLYPIINATSKDSDVTEVMW 93
DB 80 PM-----QDEGLEIWE-----QVHKVKTQEQSARNIMLVRYNKSMDPHT--LQW 127
QY 94 OPALRRGRGLQAQGYGVRIQDAGVLLYSVLFQDY-----FTMGQVYS 138
DB 128 Q-----QGCDEVG--SDGRLLHWIDQLAFDGVDPHTLNKDLRFWTAWTSTVAQISO 175
QY 139 RESGQROETLFCIRSMPSHPDRAYNSQYSAGVPHLH-----QGD 179
DB 176 PELEARLKD--NCSELLQKYPEKEKERELLRSDBPRAHVTRQPRPEGDV 221

Search completed: May 14, 1999, 21:40:00
Job time: 7994 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:30:53 ; Search time 2254.51 Seconds

(without alignments)
1417.049 Million cell updates/sec

Title: US-09-212-270-8

Perfect score: 1 tccggcgccgcggggcagga.....taaaagtctctatacccta 893

Scoring table:

Search: 808301 segs, 1788773984 residues

Database :

GenEmbl.*
1: gb_da1.*
2: gb_da2.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl1.*
9: gb_pl2.*
10: gb_pr1.*
11: gb_pr2.*
12: gb_pr3.*
13: gb_ro.*
14: gb_sy.*
15: gb_sy.*
16: gb_un.*
17: gb_vl.*
18: gb_vtg.*
19: em_da.*
20: em_fun.*
21: em_hum1.*
22: em_hum2.*
23: em_in.*
24: em_om.*
25: em_or.*
26: em_ov.*
27: em_pat.*
28: em_ph.*
29: em_pl.*
30: em_ro.*
31: em_sy.*
32: em_un.*
33: em_vl.*
34: em_vtg.*
35: em_sts.*
36: gb_da1.*
37: gb_da2.*
38: gb_pl1.*
39: gb_pl2.*
40: gb_pr1.*
41: gb_pr2.*
42: gb_pr3.*
43: gb_sts.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
--------	-------	-------	-------	--------	----	----	-------------

1	36.8	4.1	3520	3	DMU23545
2	36.8	4.1	9894	11	HS113D11
3	36.8	4.1	9894	41	HS113D11
4	36.6	4.0	3339	10	HUM8DC62
5	36.6	4.0	60904	12	AC005612
6	36.6	4.0	110000	18	HSAT11929_5
7	36.6	4.0	3339	40	HUM8DC62
8	36.6	4.0	60904	42	AC005612
9	35.6	4.0	140977	11	AC002536
10	35.6	4.0	140977	41	AC002536
11	35.4	4.0	2315	1	BFU75371
12	35.4	4.0	2315	36	BFU75371
13	35.4	3.9	1912	13	MMRARG
14	34.8	3.9	40127	11	HSF0811
15	34.8	3.9	39872	11	HSICB2046
16	34.8	3.9	40127	41	HSF0811
17	34.8	3.9	39872	41	HSICB2046
18	34.4	3.9	4018	10	HUMRPO04
19	34.4	3.9	115835	12	HS28C20
20	34.4	3.9	4018	43	MMCOL18A10
21	34.4	3.9	115835	42	HS28C20
22	34.4	3.9	115835	42	HS28C20
23	34.4	3.8	11335	1	AE000082
24	34.4	3.8	3233	1	TTARGOPER
25	34.4	3.8	6824	2	TTARGOPER
26	34.4	3.8	11335	36	AE000082
27	34.4	3.8	3233	36	TTARGOPER
28	34.4	3.8	6824	37	TTARGOPER
29	33.8	3.8	3013	4	RABTLCPEH
30	33.8	3.8	3013	27	E09934
31	33.6	3.8	3851	12	HSITGBF07
32	33.6	3.8	3851	42	HSITGBF07
33	33.4	3.7	201577	18	AC005831
34	33.4	3.7	1376	6	I01659
35	33.4	3.7	2810	6	YSLJEU2B
36	33.3	3.7	44762	11	AC002325
37	33.3	3.7	162485	18	AC004840
38	33.3	3.7	2810	6	YSLJEU2B
39	33.3	3.7	44762	41	AC002325
40	32.8	3.7	189370	1	AF010496
41	32.8	3.7	4809	1	AF011738
42	32.8	3.7	2070	1	RHBSRA
43	32.8	3.7	4998	11	AF001042
44	32.8	3.7	211072	18	HS110883
45	32.8	3.7	189370	36	AF010496

ALIGNMENTS

RESULT	1
LOCUS	DMU23545
DEFINITION	Drosophila melanogaster HK protein mRNA, complete cds.
ACCESSION	U23545
NID	9901999
KEYWORDS	
SOURCE	
ORGANISM	fruit fly.

Drosophila melanogaster
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 546)

REFERENCE
AUTHORS Choulard,S.W., Wilson,G.F., Schlimgen,A.K. and Ganetzky,B.
TITLE A potassium channel beta subunit related to the aldol keto reductase
superfamily is encoded by the Drosophila hyperkinetic locus
Proc. Natl. Acad. Sci. U.S.A. 92 (15), 6763-6767 (1995)

JOURNAL MEDLINE
95350149

REFERENCE
AUTHORS Choulard,S.W.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-1995) Scott W. Choulard, Genetics, University of
Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA

```
FEATURES
  source
    Location/Qualifiers
      1..3520
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone="HC206"
        /clone_lib="Drosophila head library--Meyerowitz"
        /chromosome="X"
        /map="9B7-8"
        /tissue_type="head"
        /dev_stage="adult"
        1..3520
        /note="Hyperkinetic"
        /gene="Hk"
        /allele="H"
        /db_xref="FlyBase:FBgn0001203"
        710..2350
        /gene="Hk"
        /codon_start=1
        /db_xref="FlyBase:FBgn0001203"
        /product="Hk protein"
        /translation="MSMALCNTNGDSNAQSTSQSOSPAATTAAPLPPLPHSHLQPL
        ESTPLILGHESGSAAPBSGGGVADGAVISTEMPTVADGAPPLPLPQSTPQ
        PLMLAPALNLTGPTTQQLMIGMAGVISTNDNNNNNNNNNNVNTSDSENPT
        IYRCRAPLASLDCMEFSGRSISLGNPALPLRHGSTPGILRYKNGKISNVG
        LGTWPFSGVSDQAEALIKLAIESGILNPDISEAHSEIETKILIORAGMRYAVI
        TKVYVSTKSEERGISRKHIIECVASLQRLQIYIDYIIHKADPMCPMEVRYAMSY
        VIOGMAWMTGRLARMSOVMETAYNRQFNITPTIVSETHMCRECECELYDEM
        NKIGYGLAMGRLSALSLSTONGDKLFLPKGSFKTSFSTWDEIIRNNAUSPQSWG
        KDRIDGRHCHRLDLALALEKLCSPQLSIAVSLKEHPVQCILLGATSAEQJHQS
        LQSLQLPLRLSSVLMLEERILENKPVPRPMISTIALR"
        3520
        /gene="Hk"
        /note="g A nucleotides: putative"
        /db_xref="FlyBase:FBgn0001203"

BASE COUNT    1054 a      873 c      897 g      696 t

ORIGIN
polyA_site
  /gene="Hk"
  /note="g A nucleotides: putative"
  /db_xref="FlyBase:FBgn0001203"

Query Match
  Best Local Similarity 47.8%; Pred. No. 3.1; Length 3520;
  Matches 107; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 52 gccatgcacacaaagctccctaccacaaagcgtgctctgcctgcctcttc 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1652 gtcattatccacaaagcggatcccatgtgtccatgatgagtggtgcgcgcacgacttac 1711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 112 ctactccacctgacatgtaagtgtgcgaacccgagagatgtcagagcagcaaatc 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1712 gtgatccacagcagcgtggtggaatgtattgtggcgacccgtcgtgatgagccagtgagatc 1771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 172 aaggaatgcgtctggaacactgtctcctcgaacacagtgcgaactgcagagatgtctc 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1772 atggagagccattatccacactgcccgctcaactgcattcagccattgttgagagatcc 1831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 aaggaatgtgctcgtcgtacatgaggaatgcacacagtgtgtcc 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1832 gagtaccattatgttctgtgcgcaaaagtgcgaactctactctgcc 1875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
HS113D11 9894 bp DNA PRI 21-AUG-1998
LOCUS Human DNA sequence from fosmid 113D11 on chromosome 22 contains
DEFINITION ESTs, Cpg island, complete sequence.
ACCESSION 283837
NID 92578046
KEYWORDS HTG; Cpg island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 9894)
```

```
AUTHORS
  Burgess,J.
  TITLE
  JOURNAL
  COMMENT
    Submitted (15-OCT-1997) E-mail enquiries: humquerry@sanger.ac.uk
    Direct Submissions
    Clone requests: clonerequest@sanger.ac.uk
    On Oct 31, 1997 this sequence version replaced gi:1772930.
    IMPORTANT: This sequence is not the entire insert of clone 113D11.
    It may be shorter because we only sequence overlapping sections
    once, or longer because we arrange for a small overlap between
    neighbouring submissions.
    This sequence has been finished according to sequence map criteria
    as follows. An attempt is made to resolve all sequencing problems,
    such as compressions and repeats, but not necessarily within known
    annotated human repeat sequence elements (e.g. Alu). Where the
    sequence is ambiguous, there is an annotation using the 'unsure'
    feature key.
    The true left end of clone 113D11 is at 1 in this sequence. The
    true left end of clone N104C4 is at 9791.
    This sequence was generated from part of bacterial clone contigs of
    human chromosome 22, constructed by the Sanger Centre Chromosome 22
    Mapping Group. Further information can be found at
    http://www.sanger.ac.uk/HR/Chr22.

FEATURES
  source
    Location/Qualifiers
      1..9894
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="22"
        /clone="113D11"
        74..139
        /note="WIR repeat: matches 327. .393 of consensus"
        808..2279
        /note="Putative Cpg island"
        833..1251
        /note="match: 5' EST AA173131 clone 611044"
        <1279..1570
        /note="match: EST D21359"
        2069..>2815
        /note="match: multiple ESTs; match: H40796 H88498"
        repeat_region
          4787..4883
          /note="MIR repeat: matches 44. .146 of consensus"
          5509..5559
          /note="3 copies of 17 mer 88 % conserved"
          5625..5745
          /note="C rich, A poor region"
          5746..5805
          /note="2 copies of 30 mer 93 % conserved"
          5815..5898
          /note="3 copies of 28 mer 87 % conserved"
          6281..6292
          /note="Single clone region."
          8013..8355
          /note="g, A rich region"
          8179..8291
          /note="Single clone region."
          8325..8332
          /note="Single clone region."
          8333..8377
          /note="False join. Repeat sequence of unknown length."
          /note="Single clone region."
          /note="Single clone region."

BASE COUNT    1937 a      2837 c      3076 g      2044 t

ORIGIN
Query Match
  Best Local Similarity 61.5%; Pred. No. 3.6; Length 9894;
  Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 502 accagaagtgtaaccttgaagatctctccaccgtctccagccctggagacagcg 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7353 atgcgcaagagagactatcaacacgctaccacagggccacagactagacagtgggc 7412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 gtggtccgtcattctgcagtgctgtgcacagtg 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7413 tgagatgggctcagctcaaccactgtgcggc 7448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

LOCUS HS113D11 9894 bp DNA PRI 21-AUG-1998

DEFINITION Human DNA sequence from fosmid 113D11 on chromosome 22 contains ESTs, Cpg island, complete sequence.

ACCESSION 283837

NID 92578046

KEYWORDS HTG; Cpg Island.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 9894)

TITLE Direct Submission

JOURNAL Submitted (15-OCT-1997) E-mail enquiries: humquerry@sanger.ac.uk

COMMENT Cloned requests: clonerequests@sanger.ac.uk

On Oct 31, 1997 this sequence version replaced gi:1772930.

IMPORTANT: This sequence is not the entire insert of clone 113D11. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The true left end of clone 113D11 is at 1 in this sequence. The true left end of clone N104C4 is at 9791.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>.

location/Qualifiers

1..9894

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="22"

/clone="113D11"

74..139

repeat_region /note="MLTID repeat: matches 327..393 of consensus"

misc_feature 808..2279 /note="Putative Cpg Island"

prim_transcript 833..1251 /note="match: 5' EST AA173131 clone 611044"

prim_transcript <1279..1570 /note="match: EST D21359"

prim_transcript 2069..2815 /note="match: multiple ESTs; match: H40796 H88498"

repeat_region 4787..4883 /note="match: multiple ESTs; match: H40796 H88498"

repeat_region 5509..5559 /note="MIR repeat: matches 44..146 of consensus"

repeat_region 5625..5745 /note="3 copies of 17 mer 88 & conserved"

repeat_region 5746..5805 /note="C rich, A poor region"

repeat_region 5815..5898 /note="2 copies of 30 mer 93 & conserved"

repeat_region 6281..6292 /note="3 copies of 28 mer 87 & conserved"

unsure

repeat_region /note="Single clone region."

repeat_region 8013..8355 /note="G, A rich region"

unsure 8179..8291 /note="Single clone region."

unsure 8325..8332 /note="Single clone region."

unsure 8333..8377 /note="false join. Repeat sequence of unknown length."

BASE COUNT 1937 a 2837 c 3076 g 2044 t

ORIGIN

Query Match 4.1%; Score 36.8; DB 41; Length 9894;

Best Local Similarity 61.5%; Pred. No. 3.6;

Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 502 accagcaagttgaacctgtggaagatctctccacgcgtccacgcttggaagcg 561

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 7353 ATCGCAAGAGGAGGACCTCATCAACACCGCTACACAGGCGCAGAGTGGCG 7412

QY 562 gtgctgcgcgtcatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 597

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 7413 TGAGATGGCTCACTCAACACCGCTGCTGCGCG 7448

RESULT 4

LOCUS HUM8DC62/c 3339 bp DNA PRI 22-AUG-1994

DEFINITION Homo sapiens (subclone H8_4_b9 from P1 35 H5 C8) DNA sequence.

ACCESSION I35661

NID 9532026

KEYWORDS MX1 region.

SOURCE Homo sapiens (library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8)) DNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 3339)

TITLE Martin,C.H., Bondoc,M.M., Chiang,A., Cloutier,T., Davis,C.A., Ericsson,C.L., Jankovic,M.A., Kim,R.J., Lee,M.T., Li,M., Mayeda,C.A., Stelert-Bi Kheir,A. and Palazolo,M.J.

JOURNAL Sequencing of the MX1 region on human chromosome 21

COMMENT Unpublished (1994)

Sequence submitted by: Human Genome Center and Drosophila Genome Center

Lawrence Berkeley Laboratory

Berkeley, CA 94720

e-mail: seqgenome.lbl.gov

This subclone overlaps H8_6_e2 and H8_3_d9.

location/Qualifiers

1..3339

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_lib="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"

BASE COUNT 749 a 969 c 875 g 746 t

ORIGIN

Query Match 4.0%; Score 36; DB 10; Length 3339;

Best Local Similarity 49.0%; Pred. No. 5.3;

Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 228 gtccaagaaatgtgcttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 287

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 392 GTACAATGATTAAAGGAGGATGATGAGGTAAAGACACAGTGAAGGCGCAGC 333

QY 288 ccggttcaaggaagacgtcggttcctccgaagtgaagcctgtcgtcgtcgtcgt 347

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 332 CCAGTACACGCACTCCAGGAGGATCCAGAGTCCCGCTTAGTCTGAGTGGAGGGA 273

QY 348 gaaccgcttcaagaggcgaactgtctacacacacagtgatgtctcgtcggaagtcgt 407

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 272 AATGTGTTCTCCAGGAGACTGCTCTGCGCTGCCAGCTGCCAAAGTCAAGATAAGCT 213

QY 408 gccaggaatttaccgg 423

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 212 TTCAGAATCTCACTG 197

RESULT 5

AC005612

LOCUS	AC005612	60904 bp	DNA	PR1	04-SEP-1998
DEFINITION	Homo sapiens chromosome 21, p1 clone LBL#8 (LBNL H8), complete sequence.				
ACCESSION	AC005612				
NID	g3540153				
KEYWORDS	HTG.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 60904) Kimmerly,M., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Mignuel,T., Miller,C., Pitluck,S., Pollard,M., Rojesteri,H., Subramanian,S. and Martin,C.H.				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 60904)				
REFERENCE	Ricke,D.O.				
AUTHORS	Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 60904)				
AUTHORS	Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojesteri,H., Subramanian,S. and Martin,C.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.				
COMMENT	Sequence submitted by: DOE Joint Genome Institute.				
FEATURES	Location/Qualifiers				
Source	1..60904 /organism="Homo sapiens" /db_xref="taxon:9606" /map="21" /clone="p1 LBL#8" /chromosome="21" /note="LBNL H8" 1..675 /rpt_family="MER25" 423..565 /rpt_family="L1" 815..4661 /rpt_family="L1" 4917..4940 /note="(ACF)8" /rpt_type=tandem /rpt_unit=ACT complement(5525..5890) /rpt_family="THER1" complement(6688..6819) /rpt_family="Alu" 7441..7741 /rpt_family="Alu" 7724..7743 /note="(A)20" /rpt_type=tandem /rpt_unit=A complement(8063..8123) /rpt_family="MERS" complement(8811..9778) /rpt_family="LTR5" complement(9970..10409) /rpt_family="MER41" complement(11266..11344) /rpt_family="LTR12" 12169..12204 /note="(GT)18" /rpt_type=tandem /rpt_unit=GT complement(13074..13520) /rpt_family="LTR7"				
repeat_region	complement(13974..14029) /rpt_family="MERS"				
repeat_region	15399..15473 /rpt_family="L1"				
misc_feature	15697..15808 /note="GRAIL 2 excellent exon, frame 0"				
repeat_region	complement(15858..16136) /rpt_family="Alu"				
misc_feature	16565..16698 /note="GRAIL 2 excellent exon, frame 1"				
repeat_region	17631..17863 /rpt_family="Alu"				
repeat_region	complement(18991..19253) /rpt_family="Alu"				
repeat_region	complement(19424..19715) /rpt_family="Alu"				
misc_feature	20241..20286 /note="GRAIL 2 excellent exon, frame 0"				
repeat_region	21081..21263 /rpt_family="MERS"				
repeat_region	21861..22013 /rpt_family="MER20"				
repeat_region	complement(23391..23728) /rpt_family="THER1"				
misc_feature	23779..23994 /note="GRAIL 2 excellent exon, frame 0"				
repeat_region	complement(24012..24087) /rpt_family="MUT1"				
misc_feature	24957..25094 /note="GRAIL 2 excellent exon, frame 1"				
repeat_region	25319..25585 /rpt_family="Alu"				
misc_feature	27635..27825 /note="GRAIL 2 excellent exon, frame 0"				
repeat_region	28141..28430 /rpt_family="Alu"				
misc_feature	28834..28966 /note="GRAIL 2 excellent exon, frame 0"				
misc_feature	29686..29929 /note="GRAIL 2 excellent exon, frame 2"				
repeat_region	29977..30000 /note="(GT)12" /rpt_type=tandem /rpt_unit=GT 31701..31844 /note="GRAIL 2 excellent exon, frame 2"				
misc_feature	complement(31932..32057) /rpt_family="MERS"				
repeat_region	32297..32320 /note="(T)24" /rpt_type=tandem /rpt_unit=T complement(32303..32508) /rpt_family="Alu"				
misc_feature	33424..33584 /note="GRAIL 2 excellent exon, frame 0"				
misc_feature	33976..34117 /note="GRAIL 2 excellent exon, frame 0"				
repeat_region	complement(34362..34558) /rpt_family="T199ge.r1"				
repeat_region	complement(34577..34696) /rpt_family="T199ge.r1"				
repeat_region	complement(35034..35111) /rpt_family="MUT1"				
repeat_region	complement(35713..36022) /rpt_family="Alu"				
repeat_region	36501..36803 /rpt_family="Alu"				
repeat_region	36909..36950 /note="(GT)21" /rpt_type=tandem /rpt_unit=GT 37127..37266				
misc_feature					

```

repeat_region /note="GRAIL 2 excellent exon, frame 2"
39543..39822
/rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
40592..40840
repeat_region /rpt_family="THE1"
complement(40971..41337)
repeat_region /rpt_family="MER4"
complement(41617..41717)
repeat_region /rpt_family="MER4"
complement(42881..43215)
repeat_region /rpt_family="Alu"
43942..44219
repeat_region /rpt_family="Alu"
complement(44724..45031)
repeat_region /rpt_family="Alu"
complement(46007..46178)
misc_feature /rpt_family="Alu"
46500..46725
/note="GRAIL 2 excellent exon, frame 2"
repeat_region /rpt_family="Alu"
47500..47781
misc_feature /note="GRAIL 2 excellent exon, frame 2"
complement(49242..49309)
repeat_region /rpt_family="THE1"
49935..50077
/rpt_family="THE1"
50080..51967
/rpt_family="MSTAR"
complement(50244..50543)
repeat_region /rpt_family="Alu"
complement(51595..51645)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
51968..52329
/rpt_family="THE1"
55080..55378
/rpt_family="Alu"
complement(56369..56511)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
complement(58621..58678)
repeat_region /note="GRAIL 2 excellent exon, frame 0"
59339..59470
/rpt_family="MER20"
complement(59793..59910)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN

Query Match 4.0%; Score 36; DB 12; Length 60904;
Best Local Similarity 49.0%; Pred. No. 7.6; Mismatches 100; Indels 0; Gaps 0;
Matches 96; Conservative 0;

QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgtgcccctgcaagccgca 287
DB 16307 GTACAATGATTAAAGGAGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCAC 16366
QY 288 ccggttcaaggaagactgggtttccagaagttaagccaatgtgcggactgtgcgtgct 347
DB 16367 CCACGTGACAGCCACTCCAGGAGTCCAGAAATCCGCTAGTGTGCTGGGTGAGGGA 16426
QY 348 gaaccgcttcgaagggccaactgtctacacaccagtgatgtctgtcggggactgct 407
DB 16427 AATCTGTTCTCCAGAGGACTGCTCGCTGCCAGCTGCCAAGTCAAGATTAAGCT 16486
QY 408 gccagatttacgg 423
DB 16487 TTCAGAAATCTCAGTG 16502

```

```

HSAJ11929_0 1 110000
HSAJ11929_1 100001 210000
HSAJ11929_2 200001 310000
HSAJ11929_3 300001 410000
HSAJ11929_4 400001 510000
HSAJ11929_5 500001 610000
HSAJ11929_6 600001 614109
Continuation (6 of 7) of HSAJ11929 from base 500001 (AJ011929 Homo sapiens *** SEQUEN

```

```

Query Match 4.0%; Score 36; DB 18; Length 110000;
Best Local Similarity 49.0%; Pred. No. 8.1; Mismatches 100; Indels 0; Gaps 0;
Matches 96; Conservative 0;

QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgtgcccctgcaagccgca 287
DB 48708 GTACAATGATTAAAGGAGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCACG 48767
QY 288 ccggttcaaggaagactgggtttccagaagttaagccaatgtgcggactgtgcgtgct 347
DB 48768 CCACGTGACAGCCACTCCAGGAGTCCAGAAATCCGCTAGTGTGCTGGGTGAGGGA 48827
QY 348 gaaccgcttcgaagggccaactgtctacacaccagtgatgtctcgtcggggactgct 407
DB 48828 AATCTGTTCTCCAGAGGACTGCTCGCTGCCAGCTGCCAAGTCAAGATTAAGCT 48887
QY 408 gccagatttacgg 423
DB 48888 TTCAGAAATCTCAGTG 48903

RESULT 7
LOCUS HUM8DC6Z/c 3339 bp DNA PRI 22-AUG-1994
DEFINITION Homo sapiens (subclone H8_4_b9 from P1 35 H5 C8) DNA sequence.
ACCESSION L35661
NID 9352026
KEYWORDS MX1 region.
SOURCE Homo sapiens (library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8)) DNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 3339)
Martin,C.H., Bondoc,M.M., Chiang,A., Cloutier,T., Davis,C.A.,
Ericsson,C.L., Jaklevic,M.A., Kim,R.J., Lee,M.T., Li,M.,
Mayeda,C.A., Steiert-Ei Kheir,A. and Palazolo,M.J.
Sequencing of the MX1 region on human chromosome 21
Unpublished (1994)
Sequence submitted by:
Human Genome Center and
Drosophila Genome Center
Lawrence Berkeley Laboratory
Berkeley, CA 94720
e-mail: seq@genome.lbl.gov
This subclone overlaps H8_6_e2 and H8_3_d9.
location/Qualifiers
1..3339
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_1lb="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"
BASE COUNT 749 a 969 c 875 g 746 t
ORIGIN

Query Match 4.0%; Score 36; DB 40; Length 3339;
Best Local Similarity 49.0%; Pred. No. 5.3; Mismatches 100; Indels 0; Gaps 0;
Matches 96; Conservative 0;

QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgtgcccctgcaagccgca 287
DB 392 GTACAATGATTAAAGGAGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCACG 333

```



```
repeat_region /note="GRAIL 2 excellent exon, frame 0"
complement(34362..34558)
repeat_region /rpt_family="Tiger1"
complement(34577..34696)
repeat_region /rpt_family="Tiger1"
complement(35034..35111)
repeat_region /rpt_family="MLT1"
complement(35713..36022)
repeat_region /rpt_family="Alu"
36501..36803
repeat_region /rpt_family="Alu"
36909..36950
repeat_region /note="(GT)21"
36943..39822
misc_feature /rpt_type=tandem
37127..37266
misc_feature /rpt_unit=GT
39543..39822
repeat_region /note="GRAIL 2 excellent exon, frame 2"
40592..40840
misc_feature /rpt_family="Alu"
40592..40840
repeat_region /note="GRAIL 2 excellent exon, frame 1"
complement(40971..41337)
repeat_region /rpt_family="THE1"
complement(41617..41717)
repeat_region /rpt_family="MER4"
complement(42881..43215)
repeat_region /rpt_family="Alu"
43942..44219
repeat_region /rpt_family="Alu"
complement(44724..45031)
repeat_region /rpt_family="Alu"
complement(46007..46178)
repeat_region /rpt_family="Alu"
46500..46725
misc_feature /note="GRAIL 2 excellent exon, frame 2"
47500..47781
repeat_region /rpt_family="Alu"
complement(49242..49309)
misc_feature /note="GRAIL 2 excellent exon, frame 2"
49935..50077
repeat_region /rpt_family="THE1"
50080..51967
repeat_region /rpt_family="MSTAR"
complement(50244..50543)
repeat_region /rpt_family="Alu"
complement(51595..51645)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
51968..52329
repeat_region /rpt_family="THE1"
55080..55378
repeat_region /rpt_family="Alu"
complement(56369..56511)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
complement(58621..58678)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
59339..59470
repeat_region /rpt_family="MER20"
complement(59793..59910)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
16756 a 14066 c 14518 g 15564 t
ORIGIN
```

```
Query Match 4.0%; Score 36; DB 42; Length 60904;
Best Local Similarity 49.0%; Pred. No. 7.6;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```

```
QY 228 gtccaaggaatgctgcgtcgtatggagatgacagtgctgcccgcagcgca 287
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16307 GTACAATGATTTAAAGGAGTATAGGCTAAGACCCACAGTGAGGCTGTGAGGCCAGC 16366
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 288 ccggtccaaggaagacgtgggttcacagaagtgttaagccacgtgtggagctgtgagctgtg 347
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 16367 CCACGTGACAGCCACTCCAGGAGATCCAGAACTCCCGCTCTAGTGTGCTGCTGAGGGA 16426
QY 348 gaaccgcttcagaagggccaactgtcacacacagtgatgtcttcgaggactgct 407
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16427 AATCTGTCTCTCCAGGAGACCTGCTGCTGCTGCCACTCCCAAGTCAGAAATTAAGT 16486
QY 408 gccaggaatttaccgg 423
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16487 TTCAGAAATCTCACTG 16502
```

```
RESULT 9
AC002536 AC002536 140977 bp DNA PRI 10-DEC-1997
LOCUS Human Chromosome 11 pac pDJ1075f20, complete sequence.
DEFINITION AC002536
NID 92673898
KEYWORDS HMG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 140977)
```

```
REFERENCE
AUTHORS Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Title: Unpublished
2 (bases 1 to 140977)
```

```
JOURNAL
REFERENCE
AUTHORS Evans,G.A., Bradbury,P., Brignac,S., Bumester,R., Burbee,D.,
Davie,J., Davies,C.J., Davis,C., English,C., Fondon,T.,
Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O.,
Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U.,
Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Title: Direct Submission
JOURNAL Submitted (11-SEP-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 140977)
```

```
REFERENCE
AUTHORS Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Title: Direct Submission
JOURNAL Submitted (10-DEC-1997) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
On Dec 10, 1997 this sequence version replaced gi:2392758.
```

```
COMMENT
FEATURES
source
1..140977
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pDJ1075f20"
/chromosome="11"
```

```
BASE COUNT 34355 a 37819 c 37235 g 31568 t
ORIGIN
```

```
Query Match 4.0%; Score 35.6; DB 11; Length 140977;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
```

```
QY 159 gcaacgagaatccaagatcgcattgaaactgtgtctctgcgaagaagtcgagctgtg 218
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 55124 GGAGGAGCAATTCACGACCCGAGGCGAGAGCCGCCCTGTTAACATCTCTGAGAGG 55183
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

OY		219	cattgaagttgtccaaagaatgctggcgcatacgtagtgaggagtgcacagtgctgccccg	278
Db	55184	CCTGCCAGATGTCCTCGAATTGATGAITACCCGTGGAAACATTTTACCCTCCTTGACGCCTCAC	55243	
OY		279	caggcgaccgccgttccaaggaaactggyggttcacaagaagttaaacgatcgtlcygcgac	336
Db	55244	CCTGCCAGATGTCCTCGAATTGATGAITACCCGTGGAAACATTTTACCCTCCTTGACGCCTCAC	55301	
RESULT	10			
LOCUS	AC002536		PRI	10-DEC-1997
DEFINITION	Human chromosome 11 pac pbdj01075f20, complete sequence.			
ACCESSION	AC002536			
NID	92673898			
KEYWORDS	HTG.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 140977)			
AUTHORS	Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignae,S., Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Geer,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.			
TITLE	Unpublished template			
JOURNAL	2 (bases 1 to 140977)			
REFERENCE	Evans,G.A., Bradbury,P., Brignae,S., Bumester,R., Burbee,D., Davis,J., Davies,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Megarthy,C., Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.			
AUTHORS	Direct Submission Submitted (11-SEP-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA 3 (bases 1 to 140977) Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignae,S., Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Geer,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.			
TITLE	Direct Submission			
JOURNAL	Submitted (10-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA On Dec 10, 1997 this sequence version replaced gi:2392758.			
COMMENT	Location/Qualifiers			
FEATURES	Source			
	1..140977	/organism="Homo sapiens"		
	/db_xref="taxon:9606"			
	/clone="PBDJ1075F20"			
	/chromosome="11"			
BASE COUNT	34355 a. 37819 c 37235 g 31568 t			
ORIGIN				
Query Match	4.0%; Score 35.6; DB 41; Length 140977;			
Best Local Similarity	50.0% ; Pred No. 11;			
Matches	89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;			
OY		159	gcagcaagaattccaaggatcgatcttgaaacgctgtcctctcaacaacagtgcgacctg	218
Db	55124	GGAAGAACAATTCACGACCGAGGGGGCGAGCAAGCGCCCTCTTAGCATCCTCGGAGGG	55183	
OY		219	cattgaagttgtccaaagaatgctggcgcatacgtagtgaggagtgcacagtgctgccccg	278
Db	55184	CCTGCCAGATGTCCTCGAATTGATGAITACCCGTGGAAACATTTTACCCTCCTTGACGCCTCAC	55243	

Db	55184	CCGGAGTGGTCCCTGGAATGATGATTCACCTGGAAACATTTTCACCCCTTCACGGCCAC	55243
Qy	279	caggccgaccggttcaagaagactgggttccagaagtgtgaagccatgtgcgac	336
Db	55244	CCTGCCACGAGAGCCACGAAACCCCTGCTGTACTGTGGGCGCCGCGCTCCCCAC	55301
RESULT	11		
LOCUS	BFU75371	2315 bp	DNA
DEFINITION	Bacteroides fragilis transposon Tn455 transposase (tnpA) gene,		07-MAY-1997
ACCESSION	complete cds, and transposase (tnpA) gene, partial cds.		
KEYWORDS	75371		
ORGANISM	92072415		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
Source			
stem_loop			
gene			
CDS			
gene			
CDS			
BASE COUNT	739 a	533 c	498 g
ORIGIN			
Query Match	4.0%	Score 35.4;	DB 1: Length 2315;
Best Local Similarity	49.2%	Pred. No. 7.5;	
Matches	93;	Conservative	0; Mismatches 96; Indels 0; Gaps 0;


```
Qy 692 gaagtagctgttagcatgtccaccagaagcttcacagaacacacgtgagacctaaag 751
|||||
Db 1320 GAACGTACCTTGTGAAACGATGAACGCTACTATACCGTTGACGAGGTGCACCAAAATCTATG 1379
Qy 752 accttagagcatcaacagctacttagaatacaagatgcagagaacagcctcttcagg 811
|||||
Db 1380 GTCTTTCGTGAGCACAACATCGCATATCGTCAAGGTGAAGCACAATGAAATAAAGG 1439
Qy 812 aatctcagaagcctccctcagaagatgctggcaagcgtgtgatatctcagaaggtaccagaagaaa 871
|||||
Db 1440 TGGGTGTGAAAAACCTGCTTTTGCCTCAGATGTGAGCGTGTCAATGCTGAAGAAGACA 1499
Qy 872 aataaagt 880
|||||
Db 1500 AATAACCGT 1508

RESULT 12
BFU75371 2315 bp DNA BCT 07-MAY-1997
LOCUS Bacteroides fragilis transposon Tn4555 transposase (tnpA) gene,
DEFINITION complete cds, and transposase (tnpB) gene, partial cds.
ACCESSION U75371
NID 92072415
KEYWORDS Bacteroides fragilis.
SOURCE Bacteroides fragilis.
ORGANISM Bacteroides fragilis.
REFERENCE 1 (bases 1 to 2315)
AUTHORS Tribble,G.D., Parker,A.C. and Smith,C.J.
TITLE The Bacteroides mobilizable transposon Tn4555 integrates by a
site-specific recombination mechanism similar to that of the
gram-positive bacterial element Tn916
JOURNAL J. Bacteriol. 179 (8), 2731-2739 (1997)
MEDLINE 97252504
REFERENCE 2 (bases 1 to 2315)
AUTHORS Tribble,G.D., Parker,A.C. and Smith,C.J.
TITLE Direct Submision
JOURNAL Submitted (18-OCT-1996) Microbiology and Immunology, East Carolina
University, Moye Blvd., Greenville, NC 27858-4354, USA
FEATURES
source 1..2315
location/Qualifiers
1..2315
/organism="Bacteroides fragilis"
/transposon="Tn4555"
/db_xref="taxon:817"
194..224
/function="transposon termini"
569..1504
/stem_loop
569..1504
/gene="tnpA"
569..1504
/gene="tnpA"
/codon_start=1
/transl_table=1
/product="transposase"
/db_xref="PID:92072417"
/translacion="KATRKCSFGCKSEYTRSGCMORYCEACQAEKRRARVNOKNNLF
KVAQPLMEIOHDEYLFESKAALMGCSROYIYKLAIGLKASRISNRAFLRRDIE
OMLEGNPYHRLPGNTSTPRKSSSLPAKRRKRESESEVDFYSGEVMALFYKO
SWLYTSAKRNHPICRIGAKNYKSKHIDEFGVANDISEITDMLTEVEVDFGMPK
TALRAYTRHKIPIKREYGRITYKSHNELRRTDLYNDERYTYEQVQITGLSSAN
ICHYKVKHIEIKIKGVKNLLRSDEYRMAERNK"
1646..25315
/gene="tnpB"
1646..25315
/gene="tnpB"
/note="integrase"
/codon_start=1
/transl_table=1
/product="transposase"
/db_xref="PID:92072416"
/translacion="MSKCTVTLRKRRIKNGTYSCLDYPGYNVTMRYTREAL
GIYIFAKPANOOERDPNARMKKAVALRNQRYEALFNENNGFFEDTKMKGDPLAYFKG
LADRNKIKMQUHYKHFQRFVNGKCTFEEDVDLCRKFMEYLLDAPQSIHTNOKLHINS
```

```
BASE COUNT 739 a 533 c 498 g 545 t
ORIGIN
Query Match 4.0%; Score 35.4; DB 36; Length 2315;
Best Local Similarity 49.2%; Pred. No. 7.5;
Matches 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
Qy 692 gaagtagctgttagcatgtccaccagaagcttcacagaacacacgtgagacctaaag 751
|||||
Db 1320 GAACGTACCTTGTGAAACGATGAACGCTACTATACCGTTGACGAGGTGCACCAAAATCTATG 1379
Qy 752 accttagagcatcaacagctacttagaatacaagatgcagagaacagcctcttcagg 811
|||||
Db 1380 GTCTTTCGTGAGCACAACATCGCATATCGTCAAGGTGAAGCACAATGAAATAAAGG 1439
Qy 812 aatctcagaagcctccctcagaagatgctggcaagcgtgtgatatctcagaaggtaccagaagaaa 871
|||||
Db 1440 TGGGTGTGAAAAACCTGCTTTTGCCTCAGATGTGAGCGTGTCAATGCTGAAGAAGACA 1499
Qy 872 aataaagt 880
|||||
Db 1500 AATAACCGT 1508

RESULT 13
MMRARG/c
LOCUS MMRARG 1912 bp RNA ROD 09-APR-1991
DEFINITION Mouse mRNA for retinoic acid receptor gamma.
ACCESSION X15848
NID 953908
KEYWORDS developmental regulation; retinoic acid receptor gamma.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthalia; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1912)
AUTHORS Zelent,A., Krust,A., Petkovich,M., Kastner,P. and Chambon,P.
TITLE Cloning of murine alpha and beta retinoic acid receptors and a
novel receptor gamma predominantly expressed in skin
JOURNAL Nature 339 (6227), 714-717 (1989)
MEDLINE 89295563
FEATURES
source 1..1912
location/Qualifiers
1..1912
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="embryonic"
/clone_lib="CDNA in lambda gt10"
1..1855
/note="Retinoic acid receptor gamma"
/evidence="experimental"
316..1693
/codon_start=1
/product="Retinoic acid receptor gamma"
/db_xref="PID:953909"
/db_xref="SWIS-PROT:P19811"
/translacion="MATKKEKLFAGALGPGSGYGGAGFPFAFPALKSGPPEMLSP
SFRGAGDPDLPEMASLSVETGTSISEENVPSPPPPRVRYKCFVCNDKSSGXYH
GVSCEGCKGFFRRSIOKNMYTCHRDNCLINKYTRNCOYCRKCFEYVMSKEYAV
RNDNRKKRKEVEEGSPDSELPQLEELITVSKAHOETPSCOLGKYTNSADH
RYOULGLMDKFSELATRCIIKIVFAKRLPFGTGLSTADOTITLKAACIDITLMRIC
TRYPEODMTMPSDGLTLNRQMNAGSGPLDIYAFARAGOLLPLEMDTETGCLSTAI
CLICGDRDLDEPERVYDKIQEPLLEALRLYARRRRSPQYMFPRMLMKITTDRCIGTSAI
GAERATILKMETPGMPPLIRMLENPEMFEDSSKPGPHRPAASSEDEAPGGQGRQ
SPQDQGP"
BASE COUNT 427 a 597 c 534 g 354 t
ORIGIN
Query Match 3.9%; Score 35; DB 13; Length 1912;
```



```
incomplete repeat"
repeat_region 7065..7353
/note="Alusx repeat: matches 292. .1 of consensus"
repeat_region 7401..7549
/note="FLAM_C repeat: matches 1. .133 of consensus"
repeat_region 7555..7854
/note="Alusg repeat: matches 1. .300 of consensus"
repeat_region 8081..8355
/note="Alusx repeat: matches 28. .298 of consensus;
incomplete repeat"
repeat_region 8384..8572
/note="Alusx repeat: matches 82. .268 of consensus;
incomplete repeat"
repeat_region 8573..8698
/note="Alub repeat: matches 273. .146 of consensus;
incomplete repeat"
repeat_region 8716..9014
/note="Alusp repeat: matches 303. .1 of consensus"
repeat_region 9037..9337
/note="Aluv repeat: matches 301. .1 of consensus"
repeat_region 9343..9635
/note="Alusx repeat: matches 294. .1 of consensus"
repeat_region 9646..9946
/note="Alusx repeat: matches 299. .1 of consensus"
repeat_region 9948..10083
/note="Aluo repeat: matches 138. .1 of consensus;
incomplete repeat"
repeat_region 10125..10425
/note="Alusx repeat: matches 302. .2 of consensus"
repeat_region 10624..10925
/note="Aluv repeat: matches 301. .1 of consensus"
repeat_region 10939..11242
/note="Alusg repeat: matches 303. .1 of consensus"
prim_transcript <11249..>12068
/note="match: multiple ESTs; match: R99642 AA507938 CL5163
C14931 W49738; match: C16651 AA659683 F22547 AA101206
AA016121; match: H39854 W52331 AA355540 AA021197 W42613;
match: AA3855141 AA017437 AA082309 AA507244 W6331; match:
AA524194 W25545 AA493910 AA583064 AA641593; match:
AA524197 AA128792 AA487370 AA487253; match: AA652667
W74358 AA599280 AA54955 AA307689; match: AA679474 W77783
AA486587 AA679458 W67984; match: AA102501 C06289 AA056113
W03710 AA150605; match: N48018 N34178 N86412 AA311933
AA071084; match: AA576022 AA580046 W05454 AA149118
AA314746; match: AA376342 AA079802 AA630792 AA487027"
gene complement(11465..11982)
/gene="BING3"
complement(<11465..11982)
CDS /gene="BING3"
/codon_start=1
/pseudo
/product="cicB2046.1"
/db_xref="PID:e1273281"
12166..12460
/note="Alusp repeat: matches 1. .296 of consensus"
repeat_region 12462..12497
/note="18 copies of 2 mer 86 & conserved"
repeat_region 12572..12802
/note="Alusx repeat: matches 1. .225 of consensus;
incomplete repeat"
repeat_region 12987..13285
/note="Aluv repeat: matches 301. .1 of consensus"
repeat_region 13288..13578
/note="Alusc repeat: matches 292. .3 of consensus"
repeat_region 13635..13929
/note="Alusg repeat: matches 295. .1 of consensus"
repeat_region 14229..14410
/note="Alusx repeat: matches 568. .382 of consensus"
repeat_region 14413..14713
/note="Alusx repeat: matches 1. .302 of consensus"
repeat_region 14728..14813
/note="MLTIE repeat: matches 387. .308 of consensus"
repeat_region 14941..15244
/note="Alusg repeat: matches 299. .1 of consensus"
repeat_region 15245..15430
/note="Aluo repeat: matches 117. .302 of consensus;
incomplete repeat"
repeat_region 15456..15564
/note="MER42c repeat: matches 1043. .1148 of consensus"
repeat_region 15583..15749
/note="FRAM repeat: matches 1. .166 of consensus"
repeat_region 15783..16044
/note="Alusx repeat: matches 34. .302 of consensus;
incomplete repeat"
repeat_region 16058..16156
/note="MER42c repeat: matches 1193. .1295 of consensus"
repeat_region 16147..16349
/note="MLTIC repeat: matches 466. .265 of consensus"
repeat_region 16350..16373
/note="12 copies of 2 mer 100 & conserved"
repeat_region 16375..16542
/note="Alusx repeat: matches 286. .124 of consensus;
incomplete repeat"
repeat_region 16543..16835
/note="Alusg repeat: matches 300. .1 of consensus"
repeat_region 16839..16971
/note="Alusg repeat: matches 133. .1 of consensus;
incomplete repeat"
repeat_region 16991..17195
/note="MLTIC repeat: matches 246. .27 of consensus"
repeat_region 17206..17448
/note="LTPA11 repeat: matches 911. .655 of consensus"
repeat_region 17471..17773
/note="Alusg repeat: matches 300. .1 of consensus"
repeat_region 17783..18080
/note="Aluv repeat: matches 301. .1 of consensus"
repeat_region 18108..18148
/note="MLTIB repeat: matches 41. .1 of consensus"
repeat_region 18168..18295
/note="MER42c repeat: matches 1414. .1538 of consensus"
repeat_region 18700..18867
/note="Aluv repeat: matches 291. .124 of consensus;
incomplete repeat"
repeat_region 18868..19164
/note="Aluv repeat: matches 297. .1 of consensus"
repeat_region 19165..19299
/note="Aluv repeat: matches 135. .1 of consensus;
incomplete repeat"
repeat_region 19376..19650
/note="Aluo repeat: matches 6. .279 of consensus;
incomplete repeat"
repeat_region 19654..19946
/note="Alusg repeat: matches 3. .294 of consensus"
repeat_region 20198..20505
/note="Alusx repeat: matches 302. .1 of consensus"
repeat_region 20529..20560
/note="16 copies of 2 mer 88 & conserved"
repeat_region 20668..20971
/note="Alusx repeat: matches 1. .302 of consensus"
repeat_region 22058..22357
/note="Alusg repeat: matches 303. .1 of consensus"
repeat_region 22475..22777
/note="Alusx repeat: matches 1. .302 of consensus"
repeat_region 22863..23083
/note="Alusg repeat: matches 296. .73 of consensus;
incomplete repeat"
repeat_region 23084..23387
/note="Alusx repeat: matches 302. .1 of consensus"
repeat_region 23398..23511
/note="Aluo repeat: matches 122. .9 of consensus;
incomplete repeat"
repeat_region 23525..23824
/note="Alusg repeat: matches 299. .1 of consensus"
repeat_region 23827..24125
/note="Alusx repeat: matches 297. .1 of consensus"
repeat_region 24128..24422
```

```
/note="AluX repeat: matches 296. .1 of consensus"
24490. .24801
/note="AluSg repeat: matches 1. .295 of consensus"
25092. .25405
/note="AluY repeat: matches 297. .1 of consensus"
25459. .25581
/note="AluO repeat: matches 3. .125 of consensus;
incomplete repeat"
25584. .25880
/note="AluSx repeat: matches 299. .1 of consensus"
25881. .26042
/note="AluO repeat: matches 115. .277 of consensus;
incomplete repeat"
26062. .26361
/note="AluSc repeat: matches 299. .1 of consensus"
26374. .26423
/note="25 copies of 2 mer 88 & conserved"
<28015. .>32406
/note="match: multiple ESTs; match: AA129657 AA386922
245843 AA405989 T79567; match: W39776 H13819 H49831 W02333
H77913 H73773; match: H34947 AA613090 AA075116 AA383223
AA493034; match: AA317661 C03211 AA085057 AA301687 W45308;
match: AA356615 W51682 AA405738 AA355087 AA129682; match:
AA355781 C00363 AA374951 AA535586 Z41479; match: H73021
T79477 AA669990 R02432 H74158 H45863; match: AA362200
AA308321 AA075005 R48034 T36090; match: AA312483 AA234490
AA644233 N73287 T30378; match: AA523436 AA12767 R47926
AA317476 AA641449; match: AA380467 AA365654 AA306929
AA305678"
29039. .32221
/gene="Daxx"
join(29039. .29245,29397. .30228,30373. .30584,30740. .30953,
31110. .31584,31745. .31967,32162. .32221)
/gene="Daxx"
/codon_start=1
/product="C1CB2046.2"
/db_xref="PID:e1273283"
/translation="MATANSIIVLDDDEDEAQAOPGSHPLPNAASPGAEAPSSSE
HAGRSSSSGGKCKYLENKLFEELIEKQTAHPKPVLPVLPVLPVLPVLPVLPVLPV
AERCNILSVLSARSRPAKIYIINELCTVLKAHSAKKNLINAPATTSNEDSGNP
P7HSLDPTNAENTASQSPRTGSRROIQLEOLALVYAEIRLOEKELJSELDPP
DSAYLOEARLKRLLRFGRLCELDKSSLTGVIVORIPYRGTRPEVNRRIERLIN
KGKPTDFPDYGLVLAVERAAAHSLGLRQOQLAAQAFVSGIRLOERHLDIYI
NGCHLTDDYRCPDPAISDPVYARLRREMSLRMSLRLEISYKVMLODKSEEGEK
KRRARLQSSHSADTPESALDSGEGSPSMAISGCGSARAEVDDDDDEDEDEEE
EEEEERATDSEEREDLEQMOBQDEDEDEDEEAAAGDGKSPSSQISNEKNL
EPEKQISRSSEGOQNGRTVSPSLISEEPLAPSSIDAENGQPELITLESBPVQL
FLELEALPLDTPSSVETDISSSRKQSEEFITVLNGAGMVSITFNGVSPFMNGD
SGPPCKSRKREKQTSGLPANSYVERORSVHEKNKCTCTLPSPSPSLASLAVADS
STRVDSHGLVTSLSCLIPPARLISOTPHSQPPRPCTCTSVATQCDPELIIYLSDD
"
repeat_region
32523. .32739
/note="Aluub repeat: matches 302. .88 of consensus;
incomplete repeat"
33739. .33907
/note="MIR repeat: matches 20. .206 of consensus"
34049. .35953
/gene="BING1"
34049. .35953
/gene="BING1"
/gene="BING1"
/product="C1CB2046.3"
/db_xref="PID:e1273284"
/db_xref="PID:g2961408"
/translation="MEPSPLSPGALPLPLSLAPPLPLPAAVVHVSPFEVTSALL
ESLMOORLOGOLDVIRVQGEFFRAHRAVLAASSPYFHDQVLLKGMTSISLSPMDP
GAFFYLASAYTGRLSMAADIVNLFVGSVLOMWHIYDKCTELLREGASATYTTT
AAATSVTVGAGVPGSGGTAVPATMGSAASHASRSRANSQSSSNYSRPRSTPS
SSQBEPAASAVSGSGRRGGGPTFPAPVVGSGGATGSKLELADLCLDDGGDGRVAV
FAGGLRRTYTPPSLMPQKHWTYVRKGNCPATPLVLPQDDLEEEEDDLVLTEDC
DDEDELGSSRVVGGGPEATLSISVRTLSEPPDGEDEOVNCESSNDFGPEEGCGP
VAGLDSGGPTSPSSYAPSHPRPLPLDMQGNQILVFPSSSSSSSSSQAPGQPGNGAE
```

```
HGAVTVGTSVSLGVPGSVGVPGGTGSGDGNKIFLCHGKAFHSKSMRDRHVNML
NLRPDPCVPCNCKFKMKHLLTEPMKTHGTLKRYEGGVCAKKFMAMDSPFRHGHERR
HRLGVGAVAPGGGTPTGSLPSPKRSRSPGVGSGSGBEASATPPSSRRVWSPPRVAKVE
MFGGSGGAN"
prim_transcript
<34893. .>35752
/gene="BING1"
/note="match: multiple ESTs; match: AA085230 AA056630
AA309450 N43919 AA468696; match: R53654 R36276 T07529
H15337 W78260 AA051601"
prim_transcript
<35824. .>36547
/note="match: multiple ESTs; match: T78477 T78559 AA359884
D82810 H34158 H15719; match: AA465000 AA062251 T07530
T03652 R53543 N34845; match: AA056605 N23868 T27481
AA455280 AA018565; match: T18609 R49563 W74632 W94380
N29958"
prim_transcript
<36919. .37233
/note="match: 5' EST AA280744 clone IMAGE:711440"
36924. .37748
/gene="tapasin"
join(36924. .36960,37101. .37271,37488. .>37748)
CDS
/gene="tapasin"
/note="continues in 297184"
/codon_start=1
/product="C1CB2046.4"
/db_xref="PID:e1273285"
/db_xref="PID:g2961409"
/translation="MKSILIAVLAIGLATAVSAGPAVTECFVENDASGKGLARPGA
LLIRGPGPPRPPLDPELYTISVDPRAGLQAAARRRPRGAPARCEMSRVPLPAS
ARKWASGLTPAQCCPRALDGMWLVISVSVLSLRLRQPPQEPVLTINAT"
7029977
/note="match: multiple ESTs; match: H95651 H95725 H95087
39072. .39100
repeat_region
/note="MER5B repeat: matches 79. .108 of consensus"
39316. .39614
repeat_region
/note="AluSp repeat: matches 303. .4 of consensus"
39623. .39872
repeat_region
/note="AluSg repeat: matches 303. .59 of consensus;
incomplete repeat"
BASE COUNT 9648 a 9448 c 10468 g 10308 t
ORIGIN
```

```
Query Match 3.9%; Score 34.8; DB 11; Length 39872;
Best Local Similarity 50.6%; Pred. No. 16;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 44 tggtagagcactgagcactcaagtcctactctacacagagcgtgctctcgctgcca 103
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36788 TGGGGAACACCGGTCGCCAGGACACCTTCACTAACACAGCGGGGATTTCACCGGCC 36847
QY 104 ttctctctactcaactcactgacatgtaagtgagtgcggaaccggagattgcaagcagc 163
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36848 CTCATGCCGCCCTTTGGAGGAAGTGAAGTGAAGAGGAGGAAGCCTTCATGCTG 36907
QY 164 aggaattcaagatgcatcgtaactgctgctctcgcaaacgctg 209
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36908 AGGAGTCCGACGCGCCTGAAGTCCCTGCTCCTCCTCGCTGTG 36953
```

Search completed: May 14, 1999, 03:31:54
Job time: 5533 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 11:42:04 ; Search time 129.53 Seconds
(without alignments)
1297.006 Million cell updates/sec

Title: US-09-212-270-8

Perfect score: 893
Sequence: 1 tccggcgccggcgggcgagga.....taaaagtgtctatacccta 893

Scoring table: IDENTITY_NUC

Searched: 240622 seqs, 94065609 residues

Database : N_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	447.2	50.1	1704	1 V33361	Nucleotide sequenc
2	447.2	50.1	1496	1 V33362	Nucleotide sequenc
3	284.8	31.9	371	1 V11422	Human secreted pro
4	69.8	7.8	181	1 V11423	Human secreted pro
5	35	3.9	1912	1 Q10388	Encodes skin-speci
6	35	3.9	1913	1 T05202	Murine retinoid ac
7	34	3.8	110000	1 V30458_2	Continuation (3 of
8	33.8	3.8	3013	1 Q94328	Rabbit telencephal
9	33	3.7	2810	1 N70218	LEU2 gene of Y. lip
10	33	3.7	2810	1 Q83184	Yarrowia lipolytic
11	31	3.5	5822	1 V62933	Human mdia Rho tar
12	31	3.5	5822	1 V62934	Human mdia Rho tar
13	30.6	3.4	3471	1 Q27167	Delta-endotoxin ge
14	30.6	3.4	3471	1 Q28940	B. thuringiensis ge
15	30.6	3.4	3471	1 Q30821	Toxin 50C. New Bac
16	30.6	3.4	3471	1 Q30853	Bt isolate Ps50C c
17	30.6	3.4	3471	1 Q51704	Bacillus thuringie
18	30.6	3.4	3471	1 T43222	Antiscarab pest to
19	30.6	3.4	674	1 V15292	Polymer clone pY12
20	30.4	3.4	7760	1 V52196	Streptococcus pneu
21	30	3.4	1878	1 V41379	Murine NF-KB recep
22	29.8	3.3	13875	1 Q20240	Murine NF-KB recep
23	29.8	3.3	13875	1 Q20240	Human centromeric
24	29.6	3.3	3366	1 Q84803	Spinocherebellar at
25	29.4	3.3	762	1 Q31724	Encodes Porcine pa
26	29.4	3.3	3946	1 T93610	Mycobacterium tube
27	29.4	3.3	49272	1 V35000	Sarcosine oxidase.
28	29.2	3.3	1710	1 Q24075	HCV envelope regio
29	29.2	3.3	565	1 Q35072	Sequence encoding
30	29	3.2	1294	1 Q45175	Phosphomycin biosy
31	29	3.2	2791	1 Q99776	Rice HYA cDNA part
32	29	3.2	1977	1 T16507	Seven transmembran
33	28.8	3.2	2098	1 Q66177	S. pneumoniae dete
34	28.8	3.2	3754	1 T28528	KSHV long unique c
35	28.8	3.2	137507	1 V19941	Human R2 seven tra
36	28.8	3.2	2098	1 V18355	Exon 2 of porcine
37	28.6	3.2	940	1 Q06394	HCV core-envelope
38	28.6	3.2	755	1 Q32449	HCV core-envelope
39	28.6	3.2	1258	1 Q32450	HCV core-envelope
40	28.6	3.2	1554	1 Q32451	HCV core-envelope
41	28.6	3.2	1258	1 Q32452	Hepatitis C virus
42	28.6	3.2	1037	1 Q58450	Hepatitis C virus
43	28.6	3.2	2033	1 Q64913	Hepatitis C virus

ALIGNMENTS

44	28.6	3.2	3366	1 Q84803	Spinocherebellar at
45	28.6	3.2	2033	1 Q86788	Hepatitis C virus
RESULT 1					
V33361					
ID	V33361	standard; cDNA to mRNA; 1704 BP.			
AC	V33361:				
DE	02-DEC-1998	(first entry)			
DE	Nucleotide sequence of human alpha-OAF065.				
KW	Human; alpha-OAF065; stroma cell; antibody; inflammatory;				
KW	cytokine-mediated disease; rheumatism; ulcerative colitis; ss.				
OS	Homo sapiens.				
FH	Key	Location/Qualifiers			
FT	CDS	45..1298			
FT		/tag= a			
FT		/product= "human alpha-OAF065 protein"			
FT	sig_peptide	45..119			
FT		/tag= b			
FT	mat_peptide	120..1295			
FT		/tag= c			
FT		/transl_except= (pos:711..713, aa= Pro)			
FT		/transl_except= (pos:714..716, aa= Arg)			
PN	MO9838304-A1.				
PD	03-SEP-1998.				
PF	26-FEB-1998: J00799.				
PR	27-FEB-1997: JP-043143.				
PA	(ONOT) ONO PHARM CO LTD.				
PI	Fukushima D, Konishi M, Tada H;				
DR	WPI: 98-481205/41.				
P	P-PSDB: W70386.				
PT	Membrane polypeptide expressed by human stroma cells, and antibodies				
PT	recognising it - for treatment of inflammatory and other				
PT	cytokine-mediated diseases.				
PS	Claim 5; Pages 31-32; 54pp; Japanese.				
CC	This is the nucleotide sequence of the human alpha-OAF065, used in				
CC	the method of the invention. The process involves the use of peptides				
CC	expressed by stroma cells, and its antibodies are used for in the				
CC	prevention and treatment of inflammatory and other cytokine-mediated				
CC	diseases such as rheumatism, ulcerative colitis.				
SQ	Sequence 1704 BP; 429 A; 426 C; 430 G; 419 T;				
Query Match	50.1%;	Score 447.2;	DB 1;	Length 1704;	
Best Local Similarity	81.6%;	Pred. No. 2e-129;			
Matches 529;	Conservative	0;	Mismatches 118;	Indels 1;	Gaps 1;
QY	32 aataaacacgttggatgagccatgcaactgaagctcctacctaaccagagcgtg	91			
DB	22 AATTAATTAATTTGTAATGAAGAAGATGGCTTTAAATGCTACAGAAACAGAAATTAATGCT	81			
QY	92 tcttcgctgcacatctctcctcctacccacgcgcatgtaagtgaattgcgaaacggag	151			
DB	82 TTTTCTACTCTTTAGTATTTACTAGGCTATTTGTCATGTAAAGTACTGTGAAACAGAG	141			
QY	152 atggcggcgagcaggaatttaagatcgaatcgtgaactgtgctcctctgaacagtgcg	211			
DB	142 ACTGTGACACAGCAAGATTCAGGATCGCTGAAACTGTGTCCCTGCACAGTGTG	201			
QY	212 gacctgcatgaggtgtgtcaagaagtgtgctcgtgctaagggaagatgacagtggtg	271			
DB	202 GGGCAGGACATGAGTGTCTCAAGATGTGCTTCGGCTATGGGAGAGATGACAGTGTG	261			
QY	272 tgcctcgtcagcgcgcagcaggttcaaggaagactgggtttccagaagtgtgaacatgtg	331			
DB	262 TGACGTGCCGCTGCACAGAGTCAAGAGACTGGGGCTTCCAGAAATGCAACCCCTGTC	321			
QY	332 cggacgtgtgctggtggaacgcgtttccagagggcgaactgctacacacacagtgatgctg	391			
DB	322 TGGACTGCGCAGTGTGTAACCGCTTTCAGAAAGCAATATGTTACAGCACAGTATGCCA	381			

Qy	392	tcctcgccggagactgcctgcgcagagattttacccggaaagaccacaactcgtgtgtgtttccaaga	451
Db	382	tcctggggggactgcctgcgcagagattttatagaagacgaacattgtggcgtttccaaaca	441
Qy	452	tgagagtgtgtccctcgcgcgagaccacactcctccactagacaaccaactgtgtaccagcaag	511
Db	442	tggagctgtgcctctgtgtggaagaccctctctctttccttgcgaacgcacattgtgtccagcaag	501
Qy	512	tgaacctgtgtaaagatctctctccacacgctctccagccctcgggacagcggtgtgtgcg	571
Db	502	tcaacctgtgtaaagatccctgcgcacagggcttcacacccacgggacacgcgcctgcctccg	561
Qy	572	tcaatcgcagtcgtctctgtgccacagctgtgtgcgcg-cctgtcatctctgtgtgtcatct	630
Db	562	ttatctgcagggcctctgtgccacacgcctcgtgcctgtcatctctgtgtgtcatctatt	621
Qy	631	gcaagagcgagttcatgtagaagaacaccagctgtaaagcttccatccc	678
Db	622	gtaacagacagtttatgtgagagaagaaacccacctgctgtctccggctcac	669

RESULT	2	
ID	V33362	
AC	V33362	
DT	02-DEC-1998	(first entry)
DE	Nucleotide sequence of human beta-OAF065.	
KW	Human; beta-OAF065; stroma cell; antibody; inflammatory;	
OS	cytokine-mediated disease; rheumatism; ulcerative colitis; ss.	
FM	Key	Location/Qualifiers
FT	CDS	45..1316
FT		/*tag= a
FT	sig_peptide	/product= "human beta-OAF065 protein"
FT		45..119
FT		/*tag= b
FT	mat_peptide	120..1313
FT		/*tag= c
FT		/transl_except= (pos:711..713, aa= Pro)
FT		/transl_except= (pos:714..716, aa= Arg)
PN	WO9838304-A1.	
PD	03-SEP-1998.	
PE	26-FEB-1998:	J00799.
PR	27-FEB-1997:	JP-043143.
PA	(ONORI) ONO PHARM CO LTD.	
PI	Fukushima D, Konishi M, Tada H;	
DR	WPI; 98-481205/41.	
DR	P-PSDB: W70387.	
PT	Membrane polypeptide expressed by human stroma cells, and antibodies	
PT	recognising it- for treatment of inflammatory and other	
PT	cytokine-mediated diseases.	
PS	Disclosure; Pages 40-41; 54p; Japanese.	
CC	This is the nucleotide sequence of the human beta-OAF065, used in	
CC	the method of the invention. The process involves the use of peptides	
CC	expressed by stroma cells, and its antibodies are used for in the	
CC	prevention and treatment of inflammatory and other cytokine-mediated	
CC	diseases such as rheumatism, ulcerative colitis.	
SO	Sequence 1496 BP; 388 A; 360 C; 372 G; 376 T;	

Query Match	50.1%;	Score 447.2;	DB 1;	Length 1496;
Best Local Similarity	81.6%;	Pred. No. 1.9e-129;		
Matches 529;	Conservative 0;	Mismatches 118;	Indels 1;	Gaps 1;

32 aataacacgatttggtagagaccatgacgtcctcaaggtcctccctacacagagcagtg 91
QY
Db 22 aataaatcattttgataagaagagggctttttaaagctctctgaaacaagagaaacgt 81
92 tcttcgcgcgcacatctctccctactccacccgcgcgcgtgaagtgaattgcgaaccggag 151
QY
Db 82 tttttacattcttttagattactagcctttttgtatgatgaaggtacattgttgaacacagag 141
141

Oy	152	atcgcagcagcagcagaattccaagatcgcacctggaacatctgtccctctgaaacagtgcg	211
Db	142	actgtgagcacacaagaatrtcagggatggctctggaaactgtgttccctgcamaaccatgttg	201
Oy	212	gacctgcatctgagttctccaaagaatctgtgctcgcgtactgsggaagatgcaacagtctg	271
Db	202	ggccagcgatcgagttgtgtctataggaaatgtggcttcggcctatggggaggaatgcacagtgtg	261
Oy	272	tgccctgcagcgcgcaccggtttcaaggaagactggygggtttccaaagtgtaaagcatgtg	331
Db	262	tgacctgcggcctctcacaggttcacaggaagactggggctttcagaaatgcamaaccctgtc	321
Oy	332	cgagctgtgcctgtgttcaaccgctttcagaaggcgcaactgtctacaacacagtgatgctg	391
Db	322	tggactgcgcagtggtgtgaacgcgttttcagaaaggaattttctacgcaccacagtgatgcc	381
Oy	392	ctctgcgggactgcctccccaagatttaccggaaagaccaactgtygtgttltcaagaca	451
Db	382	tctgggggactgtcgttccacagattttataggaagacgaactgtgtcggctttcaagaca	441
Oy	452	tggagttgtgtccctctgggaagaccactctctccctcagcaacacactgtaccagaaga	511
Db	442	tggagtgtgtcccttctgtggagaaaccttctctctcttaccgaacggcactgtgccagaa	501
Oy	512	tgaacctgtgaagaatcctcctccacgctctccagccctcgggaacaggcgtgtgctgcg	571
Db	502	tcaacctcgttaaatatgcgcgtccacggcctctccacgccacggcgacacacggcctggctgcc	561
Oy	572	tcaatctgagtgagctctggccacgctgtgtctcg-ccgtgcacatcctgtgtgtatctact	630
Db	562	ttatctgcagcgcctctggccaccgctctctggcccttgcctcatctctgtgtcatctatt	621
Oy	631	gcaaaagcagttcatcgtgagaagaacaccaagctgtaaagctccatccc	678
Db	622	gtaaagacacatttatggagaaagaaacccacagctgtgtcgtccgtcac	669

RESULT 3

ID V11422 standard; cDNA; 371 BP.
AC V11422;
DT 23-JUL-1998 (first entry)
DE Human secreted protein clone AX92_3 cDNA 5'-end.
KM Secreted protein; prevention; treatment; gene therapy; ds..
OS Homo sapiens.
PN W09801554-A2.
PD 15-JAN-1998.
PF 07-JUL-1997; U11876.
PR 09-JUL-1996; U5-677231.
PA (GENY) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI; 98-110230/10.
DR P-PSDB; W58844.
PT Secreted proteins and polynucleotides encoding them - useful to
PT prevent, treat and ameliorate medical conditions
PS Claim 15; Page 57; 93pp; English.
CC V11422-V11424 encode fragments of a novel secreted protein derived from
CC clone AX92_3 which was isolated from a human adult testes cDNA library.
CC The protein can be used to prevent, treat or ameliorate a medical
CC condition, while the polynucleotides can be used for gene therapy.
SQ Sequence 371 BP; 83 A; 96 C; 107 G; 82 T;

Query Match	31.98;	Score 284.8;	DB 1;	Length 371;
Best Local Similarity	85.28;	Pred. No. 3.1e-79;		
Matches 316; Conservative	0;	Mismatches 55;	Indels 0;	Gaps 0;

[illegible]


```

Db      61  GCTTGGGCTATGGGAGATGCACAGTGTGTACGTGCCGGGTGCACAGCTTCAAGAGAG 120
Qy      302 actggaggttcacgaagtgttaagccatgtcggactgtgcgtgtgtaacccgttcaga 361
Db      121  ACTGGGCGCTCCAAAATGCAAGCCCTGTCTGTGACTGGCGAGTGGTGAACCCCTTTCAGA 180
Qy      362  gggccaactgtctcacaccacccagtgatgtctgtctgcggggagaccgtcgtccagagatttacc 421
Db      181  AGGCAATTTGTCACCCACCACTGATGCATCTGGGGGACGTGGCCANAGATTATTATA 240
Qy      422  ggaagaccacaaactgtgtgttttcaagaacatgagtggtgtgcccgcgagagaccacctc 481
Db      241  GGAAGACANAACCTTGTCCGCTTTCAAAACATGAGAGTGTGTGCTTGTGGAACCCCTCTC 300
Qy      482  ctccctacgaacacacactgtacacgaagtgagtgacacctgtgaaagatctctccaccgtct 541
Db      301  CTCCTTACGAACCCGACCTGTGCACGAAGTCAAGTCACTCGTAAATCNCGTCCACGGCCT 360
Qy      542  ccagccctcgg 552
Db      361  CCAGCCACGG 371

```

```

RESULT 4
V11423
ID V11423 standard: cDNA; 181 BP.
AC V11423;
DE 23-JUN-1998 (first entry)
DE Human secreted protein clone AX92.3 cDNA internal fragment.
KW Secreted protein; prevention; treatment; gene therapy; ds.
OS Homo sapiens.
PN W0801554-A2.
PD 15-JAN-1998.
PE 07-JUL-1997; U11876.
PR 09-JUL-1996; US-677231.
PA (GENY ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, LaVallie ER, McCoy JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-110230/10.
DR P-PSDB; W58844.
PT Secreted proteins and polynucleotides encoding them - useful to
PT prevent, treat and ameliorate medical conditions
PS Claim 13; Page 57; 93pp; English.
CC V11422-V11424 encode fragments of a novel secreted protein derived from
CC clone AX92_3 which was isolated from a human fetal brain cDNA library.
CC The protein can be used to prevent, treat or ameliorate a medical
CC condition, while the polynucleotides can be used for gene therapy.
CC Sequence 181 BP; 43 A; 47 C; 34 G; 57 T;
SQ

```

```

Query Match      7.8%; Score 69.8; DB 1; Length 181;
Best Local Similarity 86.5%; Pred. No. 1.6e-12;
Matches 77; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

```

Qy      413  gatttaccggaagaccaaactgtgttttcaagaacatgagtggtgtgccctcgagag 472
Db      88  GATTTTATAGAGAGACGAACCTTGTGGCTTTCAGACATGAGTGTGCTTGTGAG 147
Qy      473  acccacctccctccctagaacacactgt 501
Db      148  ACCCTCTCTCTCTTACGACCGCACTCT 176

```

```

RESULT 5
O10388/c
ID O10388 standard: cDNA; 1912 BP.
AC O10388;
DE 11-APR-1991 (first entry)
DE Encodes skin-specific murine Retinoic Acid Receptor-gamma.
KW mouse; retinoic acid receptor; mRAR-gamma; embryogenesis; ss.
OS Mus musculus.
FT key Location/Qualifiers

```

```

FT      cds      316..1892
FT      /*tag= a
FT      /product= murine RAR-gamma
PN      EP-41323-A.
PD      06-FEB-1991.
PE      29-JUN-1990; 112469.
PR      30-JUN-1989; US-374690.
PR      29-MAR-1990; US-502140.
PA      (INRM ) INSERM INST SANTE.
PA      (SOUT) LES LABS SOUTIB SA.
PI      Chambon P, Kastner P, Krust A, Petkovich M, Zient A;
PI      Leroy P, Mendelsohn C, Staud A;
DR      WPI: 91-038271/06.
DR      P-PSDB; R10549.
PT      Novel human and mouse retinoic acid receptors - encode proteins
PT      used to assay for agonists and antagonists
PS      Disclosure; Fig 1; 33pp; English.
CC      An 11.5 day-old total mouse embryo lambda gt10 cDNA library was
CC      screened with labelled human RAR-alpha and -beta probes. Positive
CC      clones were isolated, mapped and sequenced. A set of clones was
CC      identified having a lesser degree of homology with either RAR-alpha
CC      or -beta than murine RAR-alpha or mRAR-beta clones, respectively.
CC      The cDNA-decoded amino acid sequence of this third set of clones
CC      was obviously related to mRAR-alpha and mRAR-beta. This new member
CC      of the mouse RAR subfamily was designated RAR-gamma. The
CC      mRAR-gamma cDNA clone was used to clone 7 different murine
CC      RAR-gamma cDNA isoforms, generated by alternative splicing of at
CC      least 7 exons. The isoforms were found to differ in their 5'-UTR
CC      sequences and in their N-terminal A region, which is known to be
CC      important for differential trans-activation by other nuclear
CC      receptors.
CC      See also R10547-8, Q10389 and Q10405-8.
SQ      Sequence 1912 BP; 427 A; 597 C; 534 G; 354 T;

```

```

Query Match      3.9%; Score 35; DB 1; Length 1912;
Best Local Similarity 47.5%; Pred. No. 0.32;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

```

Qy      146  ccgaagattgcagacgaagcaatccaagatcgatcgatcgaaactgtctcctcgaaac 205
Db      233  CCTGGCCCTGGAGAGCTCCGTCCACCTCGCTGTCTGTCCAGTGGCTTTCAGAGATC 174
Qy      206  agtgcgacctgagatgagtgatgttcacgaagatgtgcttcgctatgaggagatgcac 265
Db      173  CCTGTGGCTCTGCAAGAGATCCCTCTACATCTGGCCGACGCTGAGGTGAGGACTG 114
Qy      266  agtgtgtgcttcgacgagccgacgggttcagaagactgggtttccagaagtgtgaac 325
Db      113  GTGCAACTGCTGTGAGGAGGCGGGGGAAGAGCCGCGAGCCCGAGTCCCGGC 54
Qy      326  catgtgcgagactgtgctgtgtgaacgccttcagagg 364
Db      53  GTCGGGCGGTCTCGGGGATGAGCACCGCGGTGGGGGG 15

```

```

RESULT 6
ID T05202/c
ID T05202 standard: cDNA; 1913 BP.
AC T05202;
DE 13-JUN-1996 (first entry)
DE Murine retinoic acid receptor gamma cDNA.
KW Retinoic acid receptor; mouse; RAR; RAR-gamma; transcription factor;
KW skin; ss.
OS Mus musculus.
FT key Location/Qualifiers
FT      cds      316..1692
FT      /*tag= a
FT      /product= a
FT      misc_feature 316..498
FT      /*tag= b
FT      /note= "hRAR-gamma A region"
FT      misc_feature 499..582

```

```

FT FT /tag= C
FT FT /note= "hRAR-gamma B region"
FT FT misc_feature 583..780
FT FT /tag= d
FT FT /note= "hRAR-gamma C region"
FT FT misc_feature 781..918
FT FT /tag= e
FT FT /note= "hRAR-gamma D region"
FT FT misc_feature 919..1578
FT FT /tag= f
FT FT /note= "hRAR-gamma E region"
FT FT misc_feature 1579..1689
FT FT /tag= g
FT FT /note= "hRAR-gamma F region"

EP-683227-A1.
PD 22-NOV-1995.
PF 29-JUN-1990; 110924.
PR 30-JUN-1989; US-374690.
PR 29-MAR-1990; US-502140.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (BRIM ) INST RECH SOUTIB.
PA (BRIM ) LES LAB SOUTIB SA.
PI Chandon P, Kastner P, Krust A, Leroy P, Mendelsohn C;
PI Petkovich M, Stauba, Zelent A;
PI WPI: 95-394349/51.
PI P-PSDB: R84726.

PT New cDNA encoding the human retinoic acid receptor gamma - useful
PT for identifying agonists and antagonists of RAR-gamma and detecting
PT expression in normal and diseased tissue
PS Example 1; Fig 1; 34pp; English.
CC This sequence represents the cDNA encoding the mouse retinoic acid
CC receptor gamma (RAR-gamma). This sequence was used in the isolation of
CC the human RAR-gamma sequence. RAR-gamma is a retinoic acid inducible
CC transcription factor, and is expressed specifically in the skin where it
CC is the predominant RAR. The C and E regions of the protein are the most
CC highly conserved regions through the RAR's of a species. The C region is
CC the DNA binding domain, and the E region is the ligand binding domain.
CC The B region is also well conserved within species. The N and C terminal
CC regions of segment D are conserved, whereas the central segment of this
CC region is not. The D region is thought to act as a hinge region.
CC Regions A and F are not conserved within a given species. The A, B, D
CC and F regions are conserved across species, suggesting that they have
CC specific functions, but are not required for the RAR's specific
CC physiological roles. The protein is used to detect competitive and
CC non-competitive agonists and antagonists. This sequence can be used to
CC detect receptor expression in normal and diseased tissues and cultures.
SQ Sequence 1913 BP; 427 A; 600 C; 531 G; 355 T;

Query Match 3.9%; Score 35; DB 1; Length 1913;
Best Local Similarity 47.5%; Pred. No. 0.32;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 146 ccgagatgacgagcagcaagaaatcaagatgacatgtaactgtctcctgcaaac 205
DB 233 CCGGGCTGGGAGGCTCCGCCACCTGCTGCTGCCAGTGGCTTCCAAGATC 174
QY 206 agtgcgagcctgcatgagatgttcaagaatgtgcttcgtatggggaagatgac 265
DB 173 CCTGTGCTCTGCTCAGAGAGTCCCTCTACACTAGTGGGCGAGCTGAGAGGACTG 114
QY 266 agtgtgtccctcagcagccgacacggttcaagaagatggtttccaaagtgtgaagc 335
DB 113 GTGGAACCTGCTGAGGAGGAGGGCGGGGAAGGCCCAAGCCCGAGGTCCCGGC 54
QY 326 catgtcgagctgtgctgtgtaaacgcttcaagaag 364
DB 53 GTGGGGGCTCTCGGGGATGAGACACCCGCTGGGGGG 15

```

```

WP Sequence split into 6 fragments LOCUS V30458 Accession V30458
WP Fragment Name Begin End
WP V30458_0 1 110000
WP V30458_1 100001 210000
WP V30458_2 200001 310000
WP V30458_3 300001 410000
WP V30458_4 400001 510000
WP V30458_5 500001 534720

Query Match 3.8%; Score 34; DB 1; Length 110000;
Best Local Similarity 57.5%; Pred. No. 4.5;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 502 accagaaggtgaacctgtgaagatctctccaccgtctcagccctcggagacagcg 561
DB 31583 ACCAGAGGTGACCCACCCAGAGTCCGCTCCGCTCGAAGAGCGCTGCACCTCC 31642
QY 562 gtggtgcgcgtcatctgcatgtcgtcgcaagcagtgctgcgct 607
DB 31643 TGGCATCGCGCATGTCCATGGCCATGGCCAGCATTTGCCCGCTT 31688

RESULT 8
Q94328
ID 094328 standard; cDNA; 3013 BP.
AC 094328:
DT 16-MAY-1996 (first entry)
DE Rabbit telencephalin cDNA.
KW Japanese white rabbit; brain; telencephalon; telencephalin;
KW membrane-bound protein; learning; memory; sensory disorder; ds.
OS Oryzotylagus cuniculus.
FH Key
FT cds Location/Qualifiers
FT 125..2863
FT /tag= a
FT /transl_except= pos:1124..1126, aa:Val
FT /transl_except= pos:1814..1816, aa:Thr
FT signal_peptide 125..211
FT /tag= b
FT mat_peptide 212..2860
FT /tag= c
FT /product= telencephalin

PN J07242697-A.
PD 19-SEP-1995.
PF 02-MAR-1994; 032328.
PR 02-MAR-1994; JP-032328.
PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.
DR WPI: 95-355283/46.
DR P-PSDB: R75206.
PT DNA encoding telencephalin, a telencephalon specific membranous
PT protein - used in treatment of learning, memory and sensory
PT disorders.
PS Claim 6; Fig 2 to 4; 18pp; Japanese.
CC This sequence encodes a membrane-specific protein isolated from the
CC telencephalon of Japanese white rabbits. The protein, designated
CC telencephalin, is useful for treating diseases associated with the
CC telencephalon (e.g. learning, memory, sensory, emotional and
CC language disorders).
SQ Sequence 3013 BP; 452 A; 1032 C; 1073 G; 456 T;

Query Match 3.8%; Score 33.8; DB 1; Length 3013;
Best Local Similarity 52.5%; Pred. No. 0.95;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 478 cctccctccctagaaacacatgctacacgaagtgtaacctgtgaagatcctccacc 537
DB 1982 CCTAGTCCAGGAGACCCAGCGCTCCCTAGAGACCTGGCACCCGGGATCTAGCTTGCAC 2041
QY 538 gtctcagccctcggagacacgagtggtgctgcgtcatctgcatgtcttgccacagtg 597
DB 2042 GCAACACACCCGCTGGGCTCGGGGTCAAAGATGTCGTAGAGCGGAGATGCGCGCG 2101

```


PT mechanisms of respiratory tract disease
 PS Claim 5; Page 31-37; 54pp; Japanese.
 CC This sequence encodes a human Rho target protein, mdia (mammalian
 CC diaphanous). This protein has active type Rho protein-combining ability,
 CC has proline-combining ability, has a M.W. of 150 Kda measured by sodium
 CC dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and its
 CC gene is positioned at human chromosome 5q31.2. The nucleic acid can be
 CC used for the recombinant production of the protein. The protein is useful
 CC for the elucidation of mechanism of diseases such as respiratory tract
 CC overensitiveness, bronchial asthma, acute marrow leukaemia and marrow
 CC dysplasia syndrome.
 SQ Sequence 5822 BP; 1468 A; 1448 C; 1504 G; 1402 T;

Query Match 3.5%; Score 31; DB 1; Length 5822;
 Best Local Similarity 56.3%; Pred. No. 9.6;
 Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 550 cggagacacggcggtggtcgatcatctgagtgctctgagcaggtgctgcctgc 609
 DB 4203 CTGGCTAAGTCGGGCTCAGGCTCCCTCGACACTGCCCTTCTTGCCCTCCAGGACG 4144
 QY 610 tcacccctgtgtcatctactgcaaggagcagttcattgagaa 652
 DB 4143 TGAAGCTGTATGTGATGTTGAGAGACGACGAGCCAGAGAA 4101

RESULT 12
 V62934
 ID V62934 standard; cDNA to mRNA; 5822 BP.
 AC V62934;
 DT 15-JAN-1999 (first entry)
 DE Human mdia Rho targeting protein cDNA #2.
 KW Rho protein; mdia; mammalian diaphanous; target protein; human;
 KW Rho protein-combining; proline-combining; chromosome 5q31.2; disease;
 KW respiratory tract; overensitiveness; bronchial asthma; marrow leukaemia;
 KW marrow dysplasia syndrome; ss.
 OS Homo sapiens.
 PN J1026280-A.
 PD 06-OCT-1998.
 PF 25-MAR-1997; 090170.
 PR 25-MAR-1997; JP-090170.
 PA (KIRI) KIRIN BREWERY KK.
 DR WPI; 99-002481/01.

PT Human Rho target protein and its gene - useful for elucidation of
 PT mechanisms of respiratory tract disease
 PS Disclosure; Page 37-39; 34pp; Japanese.
 CC This sequence encodes a human Rho target protein, mdia (mammalian
 CC diaphanous). This protein has active type Rho protein-combining ability,
 CC has proline-combining ability, has a M.W. of 150 Kda measured by sodium
 CC dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and its
 CC gene is positioned at human chromosome 5q31.2. The nucleic acid can be
 CC used for the recombinant production of the protein. The protein is useful
 CC for the elucidation of mechanism of diseases such as respiratory tract
 CC overensitiveness, bronchial asthma, acute marrow leukaemia and marrow
 CC dysplasia syndrome.
 SQ Sequence 5822 BP; 1402 A; 1504 C; 1448 G; 1468 T;

027167
 ID 027167 standard; DNA; 3471 BP.
 AC 027167;
 DT 27-JAN-1993 (first entry)
 DE Delta-endotoxin gene.
 KW B.t.PS50C; Colorado potato beetle; Leptinotarsa decemlineata;
 KW Pseudomonas; expression vector; ss.
 OS Bacillus thuringiensis.
 PN EP-498537-A.
 PD 12-AUG-1992.
 PF 16-JAN-1992; 300366.
 PR 16-JAN-1991; US-642112.
 PA (MYCO) MYCOGEN CORP.
 PI Roncerada L, Payne JM, Sick AJ;
 DR WPI; 92-270497/33.
 P-P-SDB; R23997.

PT New insecticidal Bacillus thuringiensis strain NRRL B-18746 - and
 PT toxin, DNA encoding it, vectors and transformed cells, effective
 PT against Coleoptera
 PS Claim 8; Page 9-11; 16pp; English.
 CC The sequence given encodes an approx. 130 kD protein known as Bacillus
 CC thuringiensis PS50C (B.t.PS50C). This novel delta-endotoxin has been
 CC shown to be active against the Colorado potato beetle (Leptinotarsa
 CC decemlineata). This toxin gene can be transformed into suitable hosts
 CC such as Pseudomonas, which can then be applied to the environment of
 CC coleopteran insects where they will proliferate and be ingested by the
 CC insects. This sequence can be inserted into an expression vector which
 CC contains a promoter/operator region, a ribosome binding site,
 CC polyadenylation signals, etc. This will allow transcription and
 CC translation of the cells in the appropriate host.
 SQ Sequence 3471 BP; 1200 A; 567 C; 751 G; 953 T;

Query Match 3.4%; Score 30.6; DB 1; Length 3471;
 Best Local Similarity 51.9%; Pred. No. 10;
 Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 757 tagagcacaacagcctactagaatacaagatcaggaagaacagcctctcaggaatc 816
 DB 1628 TAGTACACGGGTCTCGTGTACAGAGAGATATATTAATAAACCAATCTTAGATAT 1687
 QY 817 caggagcctcctaggatgagtcgcaagcgtgtgattctcgaagctaccaggaataata 876
 DB 1688 TAGGACCTTTGGCGTTACAGTAATAGCGTCGTTATCCACAAAGATATCGTTAAGATTTC 1747

QY 877 aagttgtctatc 889
 DB 1748 GCTATGCTCTTAC 1760

RESULT 14
 028940
 ID 028940 standard; DNA; 3471 BP.
 AC 028940;
 DT 24-FEB-1993 (first entry)
 DE B. thuringiensis PS50C insecticidal protein gene.
 KW B.t.PS50C; NRRL B-18746; PMTC1658; Colorado potato beetle;
 KW coleopteran-active toxin; ss.
 OS Bacillus thuringiensis.
 FH key location/Qualifiers
 FT mat_peptide 1..3471
 FT FT
 FT FT
 FT CA2059242-A.
 PN 17-JUL-1992.
 PF 13-JAN-1992; 059242.
 PR 16-JAN-1991; US-642112.
 PR 02-JAN-1992; US-812180.
 PA (MYCO) MYCOGEN CORP.
 PI Roncerada L, Payne JM, Sick AJ;
 DR WPI; 92-323784/40.
 P-P-SDB; R27343.

PT New *Bacillus thuringiensis* isolate - used for preparing toxin,
PT DNA and transformed hosts for controlling coleoptera insect pests
PS Claim 13; Page 18-20; 33pp; English.
CC Total cellular DNA was prep'd. from B. t. PS50C. The DNA was partially
CC digested with Sau3a and 9-23kb fragments were ligated into BamHI-
CC digested lambda GEM-11. Packaged phage were plated on E.coli KW251
CC cells and screened using a radiolabelled probe. Positive plaques
CC were purified and rescreened. Single isolated purified plaques were
CC used to infect E.coli KW251 in liquid culture. DNA was isolated,
CC digested with XhoI and large fragments ligated to XhoI-cut pHTNuit1.
CC The ligation mixt. was introduced into competent E.coli NM522 cells.
CC Plasmid pMYC1638 contained a 12kb XhoI insert from which the 3471bp
CC sequence of the toxin gene was determined.
50 Sequence 3471 BP; 1202 A; 570 C; 745 G; 954 T;

This Page Blank (uspto)

C	3		67.6	404	20	AA959217	AA959217 f80d4d10.
C	4	36.4	7.6	410	22	AAS52514	AAS52514 rK12n08.s
C	5	36	4.0	399	29	A10B8A589	A10B8A589 OXG63S09.S
C	6	35.2	3.9	552	23	A10E63417	A10E63417 GH03J215.S
C	7	34.8	3.9	416	25	AA69Z849	AA69Z849 vT57h10.s
C	8	34.6	3.9	301	26	AA7443377	AA7443377 nvYl8e11.s
C	9	34.4	3.9	386	12	H92t11	H92t11 yS84f03.r1
C	10	34.4	3.9	275	27	AA829315	AA829315 od05f09..s
C	11	34.2	3.8	295	18	AA359174	AA359174 EST6618r
C	12	34.2	3.8	433	21	AA279548	AA279548 zS86D08.R
C	13	34.2	3.8	434	30	A11313155	A11313155 uP9Oe02.r
C	14	34	3.8	383	14	AA055710	AA055710 zI75e03.r
C	15	33.8	3.8	553	31	A1266650	A1266650 qm45g03.x
C	16	33.6	3.8	467	10	R02555	R02555 ye80b07.s1
C	17	33.6	3.8	585	13	W02452	W02452 Zaa47D01.r1
C	18	33.4	3.7	421	11	H06858	H06858 Yl84bD02.r1
C	19	33.4	3.7	466	11	H14764	H14764 YMz24H03.r1
C	20	33.4	3.7	497	13	N90576	N90576 zb42C02.s1
C	21	33.4	3.7	309	14	M35818	M35818 mCl4G03.r1
C	22	33.4	3.7	272	18	AA341295	AA341295 EST46907
C	23	33	3.7	487	11	H20423	H20423 yn5e612.r1
C	24	33	3.7	589	13	W37947	W37947 Zc13d03.r1
C	25	33	3.7	504	22	WL0615	WL0615 ma3yD10.r1
C	26	33	3.7	233	29	RIICC1005OA	D21918 Rice cDNA,
C	27	33	3.7	333	30	A1156639	A1156639 ue53P03..r
C	28	32.6	3.7	427	10	T86351	T86351 yD84f11..s1
C	29	32.6	3.7	506	11	H14226	H14226 yM63a10..s1
C	30	32.6	3.7	415	11	R63653	R63653 yL10D03..s1
C	31	32.4	3.6	450	11	F91903	F91903 ypQ04h10.r1
C	32	32.4	3.6	437	27	AA845505	AA845505 AK04B11.s
C	33	32.2	3.6	474	21	AA292015	AA292015 zC28C10..R
C	34	32.2	3.6	477	24	AA607232	AA607232 zu81h04.S
C	35	32.2	3.6	435	28	A1007948	A1007948 EST2023939
C	36	32.2	3.6	357	28	A1013679	A1013679 EST208355
C	37	32.2	3.6	339	30	A110U988	A110U988 EST210277
C	38	32.2	3.6	402	30	A114U948	A114U948 UT-R-BTO
C	39	32	3.6	413	14	W76336	W76336 zd6OG01.r1
C	40	32	3.6	382	29	A1037602	A1037602 ub62g10..r
C	41	31.8	3.6	440	15	AA165941	AA165941 ms50e12..r
C	42	31.8	3.6	512	30	A11819167	A11819167 us33ef10.Y
C	43	31.6	3.5	368	27	AA831737	AA831737 oc93a12..s
C	44	31.6	3.5	428	30	A115S141	A115S141 ud59ft11.r
C	45	31.4	3.5	359	10	T65828	T65828 yc1ic06.r1

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:261704

Seq primer: ETPRimer
 High quality sequence stop: 345.
 Location/Qualifiers

FEATURES

source

1..401
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTTACCAATCTGAAGTGGAGCGGCGGGAATTTTCTTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2 1] double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 /db_xref="taxon:10090"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /clone_1lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 <1..>3401
 92 a 96 c 128 g 85 t

MRNA
 BASE COUNT
 ORIGIN

Query Match 37.3%; Score 333.4; DB 13; Length 401;
 Best Local Similarity 98.3%; Pred. No. 5e-93;
 Matches 337; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 19 gacaaagggaagaataaacaagcttggtgagagccatggcactcaaggtcctactcta 78.
 |||
 DB 59 GCGGAGTAGCAGGAATAAACAGCTTGTGAGAGCCATGCGACTCAAGGCTCTACCTCTA 118
 |||
 QY 79 caaagagaggtgtcttcgtcgtccattcttcttactccacttgcatgttaagtgagt 138
 |||
 DB 119 CACAGAGCGGTCTCTTCGCTGCGCATCTCTCTACTCCACCTGGCATGTAAAGTGAGT 178
 |||
 QY 139 tgcgaacccggaatgcaagcagcaggaatcaagatgcatctggaactgttccctc 198
 |||
 DB 179 TCGGAACCGGAAATTCAGAGCAGCAGAAATTCAGGATGATCTGGAATCTGTCTCTC 238
 |||
 QY 199 tgcgaacagtgcggaacttgcatgtgagttgtccaaggaatgtgtccttcggtccta 258
 |||
 DB 239 TCGAAACAGTGGCGACTGCGCATGTGATGTCCAGGAATGTGGCTTCGGCTTGGGGAG 298
 |||
 QY 259 gatgcagcagtggttgccttcgcaagccgacacgggttcaaggaagactggggttccagaag 318
 |||
 DB 299 GATGCAACAGTGTGCTTCGCAAGCGCAGCGGATTCAGGAAGACTGGGGTTCCAGAAAG 358
 |||
 QY 319 tctgaagccatgtgcgagactgtgcctggtgtaaacgccttcaga 361
 |||
 DB 359 TGTAAAGCCATGTGCGGAGCTGTGCGCTGTGTGTAACCGCTGTCTCAGA 401

RESULT 2
 AA036247 358 bp MRNA EST 26-AUG-1995
 LOCUS mi74a03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 472300
 DEFINITION 5', mRNA sequence.
 ACCESSION AA036247
 NID 91509376

KEYWORDS

SOURCE

ORGANISM

EST.
 house mouse.
 Mus musculus

Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
 Murinae; Mus.

REFERENCE

1 (bases 1 to 358)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

TITLE

JOURNAL

COMMENT

Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:283044
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 347.
 Location/Qualifiers

FEATURES

source

1..358
 /organism="Mus musculus"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTTACCAATCTGAAGTGGAGCGGCGGGAATTTTCTTTTCTTTTCTTTT
 T 3'], double-stranded cDNA was size selected, ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."
 /db_xref="taxon:10090"
 /clone_lib="Soares mouse p3NMF19.5"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 <1..>358
 81 a 98 c 110 g 69 t

MRNA
 BASE COUNT
 ORIGIN

Query Match 35.9%; Score 321; DB 14; Length 358;
 Best Local Similarity 100.0%; Pred. No. 3.4e-89;
 Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccgggcccggggggaagcaaggggaaggaataaacaagcttggtgagagccatgca 60
 |||
 DB 38 TCCGGGCGCGCGGGGAGGAGCAAGGGGAAGTAACAGCTTGTGTGAGACCCATGGCA 97
 |||
 QY 61 ctcaaggtcctacctatacaggaagcagtgctcttcgtgcacattcttctactccac 120
 |||
 DB 98 CTCAGAGTCTTACTTACACAGAGAGCGTCTTTCGCTCCATTTCTTCTTCTTCTTCCAC 157
 |||
 QY 121 ctggcactgtaagtgagtgctgcaaaacggagatctgcaaggcagcaggaaatccaagatca 180
 |||
 DB 158 CTGGCATGTAAAGTGAAGTTCGGAACCGGAGATGTGAGGACACAGGAATTCAGAGATCA 217
 |||
 QY 181 tctggaaactgtgtctcttcgcaaacagtgcggaacttgcatgtgagttgtccaaggaatgt 240
 |||
 DB 218 TCTGGAACATGTGCTCTGCAACAGTCCGAGACCTGCGCATGAGATTGTCCAGGAAATGT 277
 |||
 QY 241 ggttcggtactaggagagatgcacagtggtgcttcgagggcgacacgggttcaagga 300
 |||

Db 278 GGCTTGGGCTATGGGAGATGACAGTGTGCTGCCCTGACGGCCGACCGGTTCAAGAA 337

QY 301 gactcgggggttcacagaagtct 321
 |||
 Db 338 GACTGGGGGTTTCCAGAAAGTGT 358

RESULT 3

AA495217 404 bp mRNA EST 27-JUN-1997
 LOCUS fa04d10.r1 zebrafish ICRFzfls Danio rerio cDNA clone 10E17 5', mRNA
 DEFINITION sequence.
 ACCESSION AA495217
 NID g2225645
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Rasbora; Danio.

REFERENCE 1 (bases 1 to 404)
 AUTHORS Clark,M., Lehrach,H., Johnson,S., Marra,M., Eddy,S., Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Martin,J., Moore,B., Schellenberg,K., Seepoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 TITLE WashU Zebrafish EST Project
 JOURNAL Unpublished (1997)
 COMMENT

Contact: Steve Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Steve Johnson lab internal ID - P1_332 NOTE - For this library, the CLONE id field represents a position identifier on the original cDNA library preparation plate. cDNA library preparation: Matthew Clark. cDNA library Arrayed by: Matthew Clark. DNA sequencing by: Washington University Genome Sequencing Center clone distribution: Genome Systems, St. Louis, and Max Planck Institut fuer Molekulare Genetik, Berlin Tel +49 30 84 13 1235
 Seq primer: T7 ET from Amersham
 High quality sequence stop: 320.
 Location/Qualifiers
 1..404
 /organism="Danio rerio"
 /note="vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5'pgactcgttctagatccgacgcccctcttttttttttttt3'], on mRNA from pooled 26 somite zebrafish embryos; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin) and was not biochemically normalised. 70,000 clones from this library were arrayed on high density filters and subsequently screened by oligonucleotide hybridization fingerprinting to identify unique or minimally redundant clones for more intensive analysis."
 /db_xref="taxon:7955"
 /clone_id="10E17"
 /clone_lib="zebrafish ICRFzfls"
 /sex="mixed"
 /tissue_type="pooled 26-somite embryos"
 /lab_host="XLI-blue MRF"
 <1..>404

BASE COUNT
 mRNA
 ORIGIN
 80 a 98 c 125 g 101 t

Query Match 7.6%; Score 67.6; DB 20; Length 404;
 Best Local Similarity 81.8%; Pred. No. 1.2e-10;
 Matches 90; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 550 cgggacacggcggtcgtccgcctcagtcgtcgtcgcacggtcgtcgc-ctg 608
 |||
 Db 1 CGAGCATGCTGCTGGCTGCTGCTATCTGCAGTCTCTGGCCACTGCTCTGCGCCTC 60

QY 609 ctcatcctgtgtcctatctactgcaagagcagtcacatgagaaagaaac 658
 |||
 Db 61 TTCAATCCTCTGTGATCTACTGTAAAGACAGCTGCTGGAGAAAGAAC 110

RESULT 4
 AA552514/c 404 bp mRNA EST 05-SEP-1997
 LOCUS nk12h08.s1 NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1013343
 DEFINITION similar to TR:G757872 G757872 ENV MNNA.; mRNA sequence.
 ACCESSION AA552514
 NID g2322768
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 404)
 AUTHORS NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 1376 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 322.
 Location/Qualifiers
 1..404
 /organism="Homo sapiens"
 /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3 adaptor sequence: 5' CTCGATCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."
 /db_xref="taxon:9606"
 /clone_id="IMAGE:1013343"
 /clone_lib="NCI CGAP Co2"
 /tissue_type="tumor"
 /lab_host="SOLR (kanamycin resistant)"
 <1..>404

BASE COUNT
 mRNA
 ORIGIN
 130 a 74 c 80 g 120 t

Query Match 4.1%; Score 36.4; DB 22; Length 404;
 Best Local Similarity 51.2%; Pred. No. 0.58;
 Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 571 gtcatcgcagtgctcgcgcacaggtcgtcgtccgcgtcatcctgtgtcctact 630
 |||
 Db 270 GCCAGGTCTAATCTATTTCACAAATAGCTTTGAGCTGCTTGTGCTCAATTATA 211

QY	631	gcaagagcgacgtctatgtagaagaaccacgctgtgaagctccacctctgtctcactg	690
Db	210	GGAGCAGATTATTATTTGGTAATATTAAAGAGCAGAAACAGACTGTGAATGTGTCCACAG	151
QY	691	tgaagtgcgttgttagcatgtcaccccaagaaggtcttcaagacac	736
Db	150	TGATTACATCCAGGCATTATTGCCAGCCAGATTATTAATATGTC	105
RESULT	5		
AI084589	AI084589	399 bp	EST
LOCUS	ox3609.s1 Soares_NHMPU_S1	Homo sapiens cDNA clone IMAGE:1661032	17-Aug-1998
DEFINITION	3', mRNA sequence.		
ACCESSION	AI084589		
NID	93423012		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euteria;		
AUTHORS	Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 399)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
FEATURES	Seq primer: -40ml3 fwd. ER from Amer sham.		
source	Location/Qualifiers		
	1. .399		
	/organism="Homo sapiens"		
	/note="Organ: mixed (see below); Vector: p7773D-Pac		
	(Pharmacia) with a modified polylinker; Site_1: Not I;		
	Site_2: Eco RI; Equal amounts of plasmid DNA from three		
	normalized libraries (melanocyte 2NBH4, pregnant uterus		
	2NBH9, and fetal heart 2NBH19w) were mixed, and ss circles		
	were made in vitro. Following HAP purification, this DNA		
	was used as tracer in a subtractive hybridization		
	reaction. The driver was PCR-amplified cDNAs from pools of		
	5,000 clones made from the same 3 libraries. The pools		
	consisted of I.M.A.G.E. clones 260232-265223,		
	340488-345479, and 484488-489479."		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1661032"		
	/clone_1lb="Soares_NHMPU_S1"		
	/tissue_type="Pooled human melanocyte, fetal heart, and		
	pregnant uterus"		
	/lab_host="DH10B"		
BASE COUNT	112 a 92 c 77 g 118 t		
ORIGIN			
Query Match	4.0%; Score 36; DB 29; Length 399;		
Best Local Similarity	51.1%; Pred. No. 0.76;		
Matches	81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;		
QY	613	tccctgtgtcactactcaagagcgagctcatgtgagaagaaccacgctgtaagctcc	672
Db	6	TCCTGGAGATATTCATTCATTAAGAAAGCTTTAAGAGAGAGAAAACCCACACACAT	65
QY	673	catccctctgtctcactgtgaaagtgcgtgttaagcatgttaccaccaagagttccaaga	732
Db	66	CAGGCTCTTATTCCTCCCTGCTGCTGTTCTTTACAGTTCTTGACCTCAGAGTTTGAAAA	125
QY	733	caccctgcgtcgaactaagacctttagagcatcaac	768
Db	126	GTACAGAGTCTTACCTTCATCTCTAGGAAGAGAGAAC	161

RESULT 6 A1063417 552 bp mRNA EST 24-NOV-1998
LOCUS GH03215.5prine GH Drosophila melanogaster head POT2 Drosophila
DEFINITION melanogaster cDNA clone GH03215 Sprime similar to U23545: HK
Fgpn001203 PID:902000 SPTREMBL:Q24052, mRNA sequence.
ACCESSION A1063417
NTD g3339361
KEYWORDS EST.
SOURCE fruit fly,
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachytera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 552)
AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
Brokslein,P., Lewis,S. and Rubin,G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL unpublished (1997)

COMMENT

Contact: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 32 row: B column: 340.
High quality sequence stop: 540.

FEATURES
source
location/Qualifiers
1..552
/organism="Drosophila melanogaster"
/note="Organ: head; Vector: POT2; Site_1: EcoRI; Site_2:
XhoI; Sized fractionated CDNs were directly ligated into
POT2. Plasmid cDNA library."
/db_xref="taxon:7227"
/cd_clone="GH03215"
/cloned_by="GH Drosophila melanogaster head POT2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"

BASE COUNT ORIGIN	114 a 168 c 169 g 101 t
--------------------------	--

```

Query Match      3.9%   Score 35.2; DB 23; Length 552;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches    88; Conservative     0; Mismatches    88; Indels     0; Gaps       0

```

OY	100	gccatcttcttactacacctgtgcattgaagtgtgcgaaccggagattcagg	159
Db	6	GCCATGAGCTACTAGTGTATTACAGCAGGGCTGGCGAATTTATTTGGGGCACCGTCATGAGGC	65
OY	160	cagcagaactcaagaatgatcttgaaaactgttcctctcgcaaacagtycggaactgyc	219
Db	66	CAGGTGAATCATGATGAGGCCCTATTACAACATGCCGCCAGTTCAACTGCATCAGCCCAT	125
OY	220	atggagttccaaagaaatgttgtcttcgcctaagtgggaagatgacaagtygtgccc	275
Db	126	GTGAGGCACTCCGAGTACATATGTTCTTGCCGCAAAAGTGCAACTTACCTGCC	181

RESULT 7 AA692849 /c 416 bp mRNA EST 16-DEC-1997
LOCUS vrs7h10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
DEFINITION I124803 5' similar to TR:Q14136 Q14136 KIAA0122 PROTEIN ;, mRNA
sequence.
ACCESSION AA692849
NTD g2693787
KEYWORDS EST.
SOURCE house mouse,
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 416)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:614139
Possible reversed clone; similarity on wrong strand.
Location/Qualifiers
1. 416
/organism="Mus musculus"
/strain="B6D2 F1/7"
/notes="Organ: embryo; Vector: pBluescribe (modified); Site: 1: MluI; Site 2: SalI; Cloned unidirectionally from mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI(dT): 5'-CGGTGACCGTCGACGCTTTTCTTTT-3'. CDNAS were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB). Average insert size: 1.2 Kb."
/db_xref="taxon:10090"
/clone="1124803"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"

BASE COUNT 110 a 96 c 134 g 76 t
ORIGIN

Query Match 3.9%; Score 34.8; DB 25; Length 416;
Best Local Similarity 55.9%; Pred. No. 1.8;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 368 actgtcacacacacagtgctgtctgtcgaggagcgtgcagagattttaccggaaga 427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 177 ACTTCTACGCCACCCCTGTGCTGTCACGATGATCTGCCGCTGGATGATACCTCTGC 118

QY 428 ccaactgtgtgttttccaagacatgatgtgtgtccctcgagagaccactctcc 485
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 117 AGAGACTCTCTCGCGCGCGGCTGTGAGAGGAGGCTCTGTCCGACACCACTCTTC 60

RESULT 8
AA743377 301 bp mRNA EST 27-JAN-1998
LOCUS ny18e11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1272140,
DEFINITION mRNA sequence.
ACCESSION AA743377
NID 92782883
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 301)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bdtp/image/image.html

FEATURES
source
Insert Length: 520 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 290.
Location/Qualifiers
1. 301
/organism="Homo sapiens"
/notes="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CDBR). CDNA synthesis was primed with a Not I - Oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTTTTCTTTT-3']
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTTTTCTTTT-3']
(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1272140"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 73 a 69 c 78 g 81 t
ORIGIN

Query Match 3.9%; Score 34.6; DB 26; Length 301;
Best Local Similarity 58.1%; Pred. No. 1.9;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 552 ggaacaggggtgtgtgcgtcatctgtcagtcctgtgcacaggtgtcgtcctgc 611
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 111 GGCTGCCGCGCGGATGGCTGTCAGAGCAGAGCGAGCGCTGTGTGCTCCGCTTTC 170

QY 612 atccctgtgtcatctactcaagaagagcagttcacaggagaaga 656
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 171 ACCCTTCCCTCTCTCTTCTGCTCAGAGGAGTGATCGCAGAGAA 215

RESULT 9
H92111 386 bp mRNA EST 29-NOV-1995
LOCUS y844f03.r1 Homo sapiens cDNA clone 221501 5'.
DEFINITION H92111
ACCESSION H92111
NID 91087689
KEYWORDS EST.
SOURCE human clone-221501 primer-M13RPI library-Soares retina N2B4HR vector-p773D (Pharmacia) with a modified polylinker host-DH10B (ampicillin resistant) RsiteI-Not I Rsite2-Eco RI 1st strand cDNA was primed with a Not I - Oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTTTTCTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian male and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL ; contact the
IMAG Consortium (info@image.lnl.gov) for further information.

Contact: Maria M/Mouse Est Project
WashU-NHMI Mouse Est Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact
IMAG Consortium (infoimage.jhl.gov) for further info.

This Page Blank (uspto)

Accession	U23545
NC_009019.9	
Keywords	
Source	
Organism	<i>Drosophila melanogaster</i>
Reference	Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda; Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephydrozoa; Drosophilidae; Drosophila.
Authors	1 (bases 1 to 546) Chouinard,S.W., Wilson,G.F., Schlimgen,A.K. and Ganetzky,B.
Title	A potassium channel beta subunit related to the aldo-keto reductase superfamily is encoded by the <i>Drosophila</i> hyperknetic locus
Journal	Proc. Natl. Acad. Sci. U.S.A. 92 (15), 6763-6767 (1995)
Medline	95350149
Reference	2 (bases 1 to 3520) Chouinard,S.W.
Authors	Direct Submission
Journal	Submitted (28-MAR-1995) Scott W. Chouinard, Genetics, University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA
Features	Location/Qualifiers
Source	1..3520
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone="HC206"
	/clone_lib="Drosophila head library--Meyerowitz"
	/chromosome="X"
	/map="9B7-8"
	/issue_type="head"
	/dev_stage="adult"
	1..3520
gene	/note="Hyperknetic"
	/gene="Hk"
	/allele=""
	/db_xref="FlyBase:FBgn0001203"
	710..2350
CDS	/gene="Hk"
	/codon_start=1
	/db_xref="FlyBase:FBgn0001203"
	/product="Hk protein"
	/db_xref="PID:g902000"
	/translation="MSMAICNLNGDSGSAOSTSGOSPAPATAAPLILPHSHSTLDP ESTPILIGHEDSGSAPBEGSGGCVADGATVSIEMPTVADDAVPLPLPPOSTPQ PLMALPANLNLITGPTTQQLMTIGNGAMGYIPSTNDNNNNNNNDTSDENPTQ ILYRRLAPIASLIDMEERSGRSISGSNAPLPLRSGSPPTGLRYKLNKSGGLRISNG LGTPVSPGVSDQAAILKLAIESGINFIDISEASETEIGKILQAGMKRTAYIV TTKYVTSSEERGLSRKHITIECVASLRLOLOIDYIVIIHKRDKCPMEYPAWY VIOGKAMNCTLASOVETNEATNRCNPICITPYEOSYHRCFEKEKELIPEMY NKTRIVGIMAWGPTSMALSDTONGDKLPLPGSEFTKFSWTEDEINNALSPGWSG KDRIDGRHRCIDRLDALAALIKGCSPTOLISMSLKHBYOCLLIGASBOLHOS LOSLOLPLRSSSYMLELERLENKPRPPIWISLALR"
	3520
polyA_site	/gene="Hk"
	/note="g A nucleotides; putative"
	/db_xref="FlyBase:FBgn0001203"
BASE COUNT	1054 a 873 c 897 g 696 t
ORIGIN	
Query Match	5.98; Score 36.8; DB 3; Length 3520;
Best Local Similarity	47.88; Pred. NO. 1.2;
Matches 107; Conservative	0; Mismatches 117; Indels 0; Gaps 0;
Db	113 gccatgagcaccacaaagctcctactacacagagcagtgctctgctgcacatctctc 172
Db	1652 GTCATCATCCACACAGCGCGGATCCATGTGTCTCCCATGGAGGTGGTCCGCCATGACGTAC 1711
Db	173 ctactccacctgagcagtgaatgaagtgcagaaacggagagatgtcagagcagcaggaattc 232
Db	1712 GTGATACAGACAGGGCTGGGCAATGATTATGGGGCACCGCTCATGTGAGCCAGGTGAGATC 1771
Db	233 aagatgatctggagaaactgctctctctgcacaaacagtcggaactctgcatgtgagttgtcc 292
Db	1772 ATGGAGGCGCTATACCACTGCGCGAGTTCATCTGCATCAGCGCCCATGTGTGAGAGCATGCC 1831

[illegible]

QY 611 aacaacacaaa 621
 Db 88438 AAAAAAAAAA 88428

RESULT 5
 HUM8DC6Z/C 3339 bp DNA PRI 22-AUG-1994
 LOCUS Homo sapiens (subclone H8 4_b9 from P1 35 H5 C8) DNA sequence.
 DEFINITION L35661
 ACCESSION 9532026
 NID
 KEYWORDS MXI region.
 SOURCE Homo sapiens (library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8)) DNA.

ORGANISM Homo sapiens
 Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata;

REFERENCE 1 (bases 1 to 3339)
 Martin, C.H., Bondoc, M.M., Chiang, A., Cloutier, T., Davis, C.A.,
 Ericsson, C.L., Jaklevic, M.A., Kim, R.J., Lee, M.T., Li, M.,
 Mayeda, C.A., Steiert, E.I., Kheir, A. and Palazzolo, M.J.
 Sequencing of the MXI region on human chromosome 21

TITLE Unpublished (1994)
 JOURNAL
 COMMENT Sequence submitted by:
 Human Genome Center and
 Drosophila Genome Center
 Lawrence Berkeley Laboratory
 Berkeley, CA 94720
 e-mail: seqgenome.lbl.gov
 This subclone overlaps H8 6_e2 and H8 3_b9.

FEATURES
 source
 1. 3339
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_lib="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"

BASE COUNT 749 a 969 c 875 g 746 t
 ORIGIN

Query Match 5.8%; Score 36; DB 10; Length 3339;
 Best Local Similarity 49.0%; Pred. No. 2.2;
 Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 289 gtccaaaggaatgagctggtgcatgagggaggtgacagtggtgcccgcagggcgca 348
 Db 392 GTACATGATTAAGGAGGATGATAGGTAGGACCCACAGTGGAGGCTCTGGAGGCGAC 333

QY 349 ccggttcaaggaagactgggtttccagaagtgtaaagcactgtgcggaactgtgcgtgt 408
 Db 332 CCACTGACACGACCTCCAGGAGGATCCAGAGTCCCGCTCTAGTGTGGTGGAGGCGA 273

QY 409 gaacgccttcaagggccacactgctacacaccagtgtgctgtctgctggggaactgtc 468
 Db 272 AATCTGTCTCCAGGAGGACCTGCTCTGGGCTGCCAGCTCCAAAGTCAGAGTAAGCT 213

QY 469 gccagatttaccgg 484
 Db 212 TTCAGAAATCTCAGCTG 197

RESULT 6
 AC005612 60904 bp DNA PRI 04-SEP-1998
 LOCUS Homo sapiens chromosome 21, P1 clone LBL#8 (UBNL H8), complete
 DEFINITION sequence.
 ACCESSION AC005612
 NID 93540153
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 60904)
 AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,
 Kadner, K., Miguel, T., Miller, C., Pitluck, S., Pollard, M.,
 Rojeski, H., Sudramanian, S. and Martin, C.H.

TITLE Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 60904)
 AUTHORS Ricker, D.O.
 TITLE Large Scale Sequence Analysis and Annotation with the Sequence
 Comparison Analysis (SCAN) System

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 60904)
 AUTHORS Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M.,
 Davis, C.A., Kadner, K., Miguel, T., Pitluck, S., Pollard, M.,
 Rojeski, H., Sudramanian, S. and Martin, C.H.

TITLE Direct Submission
 JOURNAL Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome
 Institute, Lawrence Berkeley National Laboratory, MS 74-157,
 Berkeley, CA 94720, U.S.A.
 COMMENT Sequence submitted by:
 DOE Joint Genome Institute.

FEATURES
 source
 1. 60904
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="21"
 /clone="P1 LBL#8"
 /chromosome="21"
 /note="UNL H8"

repeat_region

/rpt_family="MER25"

repeat_region

/rpt_family="L1"

repeat_region

/rpt_family="L1"

repeat_region

/note="(ACT)8"

repeat_region

/rpt_type=tandem

repeat_region

/rpt_type=tandem

repeat_region

/rpt_family="MER5"

repeat_region

/rpt_family="L1"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```
repeat_region 17631..17863
/rpt_family="Alu"
repeat_region complement(18991..19253)
/rpt_family="Alu"
repeat_region complement(19424..19715)
/rpt_family="Alu"
misc_feature 20241..20286
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 21081..21263
/rpt_family="MER5"
repeat_region 21861..22013
/rpt_family="MER20"
repeat_region complement(23391..23728)
/rpt_family="THE1"
misc_feature 23779..23994
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(24012..24087)
/rpt_family="MLT1"
misc_feature 24957..25094
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 25319..25585
/rpt_family="Alu"
misc_feature 27635..27825
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 28141..28430
/rpt_family="Alu"
misc_feature 28834..28966
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 29686..29929
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 29977..30000
/note="(GT)12"
/rpt_type=tandem
/rpt_unit=GT
/rpt_unit=GT
31701..31844
/note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(31932..32057)
/rpt_family="MER5"
repeat_region 32297..32320
/note="(T)24"
/rpt_type=tandem
/rpt_unit=T
repeat_region complement(32303..32508)
/rpt_family="Alu"
misc_feature 33424..33584
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 33976..34117
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(34362..34558)
/rpt_family="T199er1"
repeat_region complement(34577..34696)
/rpt_family="T199er1"
repeat_region complement(35034..35111)
/rpt_family="MLT1"
repeat_region complement(35713..36022)
/rpt_family="Alu"
repeat_region 36501..36803
/rpt_family="Alu"
repeat_region 36909..36950
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
/rpt_unit=GT
37127..37266
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 39543..39832
/rpt_family="Alu"
misc_feature 40592..40840
/note="GRAIL 2 excellent exon, frame 1"
repeat_region complement(40971..41337)
/rpt_family="THE1"
repeat_region complement(41617..41717)
/rpt_family="MER4"
repeat_region complement(42881..43215)
```

```
repeat_region /rpt_family="Alu"
43942..44219
/rpt_family="Alu"
repeat_region complement(44724..45031)
/rpt_family="Alu"
repeat_region complement(46007..46178)
/rpt_family="Alu"
misc_feature 46500..46725
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 47500..47781
/rpt_family="Alu"
misc_feature complement(49242..49309)
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 49335..50077
/rpt_family="THE1"
repeat_region 50080..51967
/rpt_family="MSTAR"
repeat_region complement(50244..50543)
/rpt_family="Alu"
misc_feature complement(51595..51645)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 51968..52329
/rpt_family="THE1"
repeat_region 55080..55378
/rpt_family="Alu"
misc_feature complement(56369..56511)
/note="GRAIL 2 excellent exon, frame 0"
misc_feature complement(58621..58678)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 59339..59470
/rpt_family="MER20"
misc_feature complement(59793..59910)
/note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN
```

Query Match 5.8%; Score 36; DB 12; Length 60904;
Best Local Similarity 49.0%; Pred. No. 3;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
QY 289 gtccaaggaatgtgcttgcgcctatgtggagagatgcacagtgtgtccctgcagccgca 348
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16307 GTACAAATGATTAAGGAGATGATGAAGGACCCACAGTGGAGGCGCAGC 16366

QY 349 ccggtccaaggaagactgtggttcacgaagttaagccatgtgcggaactgtgcctgt 408
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16367 CCACGTACACGCCACTCCAGGAGACTCCAGAGTCCCGCTCTAGTGTGGTGTGAGGGA 16426

QY 409 gaaccgcttcagaggccaactgtctcacaccagtgatgtctgtcgcgggaactgtcct 468
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16427 AATCTGTTCCTCCAGGAGACTGTCTCGGCTGCCAGCTGCCAAGATCGAATAAGCT 16486

QY 469 gccaggaatttaccgg 484
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16487 TTCAGAAATCTCACTG 16502
```

RESULT 7
HSAJ11929_5
WPCOMMENT
Sequence split into 7 fragments LOCUS HSAJ11929 Accession AJ011929
Fragment Name Begin End
HSAJ11929_1 100001 210000
HSAJ11929_2 200001 310000
HSAJ11929_3 300001 410000
HSAJ11929_4 400001 510000
HSAJ11929_5 500001 610000
HSAJ11929_6 600001 614109
Continuation (6 of 7) of HSAJ11929 from base 500001 (AJ011929 Homo sapiens *** SEQUEN

Query Match	5.8%	Score 36:	DB 18:	Length 110000:
Best Local Similarity	49.0%	Pred. No. 3.2:		
Matches 96:	Conservative	0:	Mismatches 100:	Indels 0:
				Gaps 0:
QY 289	gtccaaagaatgtgcttcgcgtatgctatgctggggaagatgcacagtgtgtgccttcaggccgca	348		
Db 48708	GTACAAATATTAAAGGAGGTATATAGGAGTAAGACCCACACAGCTGGAGGCTTGGAGGCCAC	48767		
QY 349	ccggttcagaagaagactgtgggtttccagaagtgttaagccatgtgcggactgtgtcgtgt	408		
Db 48768	CCACTGACAGCCACTCCAGGAGGTCCAGAAAGTCCCGCTTAAGTGTGGTGTGGAGGGA	48827		
QY 409	gaacgcttcagaagggccaactgtctacacaccacagtgatgtgtctgtcgggagctgct	468		
Db 48828	AATCGTTCCTCCAGGAGCTGTGCTCGGCTGCCACGCTGCCAAAGTCAAGAAATAAGCT	48887		
QY 469	gccaaagtattacagg 484			
Db 48888	TTCACAATCTCACATG 48903			
RESULT 8	HUM8DC6/c			
LOCUS	HUM8DC6Z	3339 bp	DNA	PRI 22-AUG-1994
DEFINITION	Homo sapiens (subclone H8 4_b9 from P1 35 H5 C8)			DNA sequence.
ACCESSION	L3561			
NID	g532026			
KEYWORDS	MXI region.			
SOURCE	Homo sapiens (library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8))			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 3339)			
AUTHORS	Martin,C.H., Bondoc,M.M., Chiang,A., Cloutier,T., Davis,C.A., Mayeda,C.A., Stelert-El Kheir,A., Kim,R.J., Lee,M.T., Li,M., Erickson,C.L., Jaklevic,M.A., and Palazzolo,M.O.			
TITLE	Sequencing of the MXI region on human chromosome 21			
JOURNAL	Unpublished (1994)			
COMMENT	Sequence submitted by: Human Genome Center and Drosophila Genome Center Lawrence Berkeley Laboratory Berkeley, CA 94720 e-mail: seqgenome.lbl.gov This subclone overlaps H8 6_e2 and H8 3_d9.			
FEATURES	location/Qualifiers			
source	1..3339			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/feature_id="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"			
BASE COUNT	749 a 969 c 875 g 746 t			
ORIGIN				
Query Match	5.8%	Score 36:	DB 40:	Length 3339:
Best Local Similarity	49.0%	Pred. No. 2.2:		
Matches 96:	Conservative	0:	Mismatches 100:	Indels 0:
				Gaps 0:
QY 289	gtccaaagaatgtgcttcgcgtatgctatgctggggaagatgcacagtgtgtgccttcaggccgca	348		
Db 392	GTACAAATATTAAAGGAGGTATATAGGAGTAAGACCCACACAGCTGGAGGCTTGGAGGCCAC	333		
QY 349	ccggttcagaagaagactgtgggtttccagaagtgttaagccatgtgcggactgtgtcgtgt	408		
Db 332	CCACTGACAGCCACTCCAGGAGGTCCAGAAAGTCCCGCTTAAGTGTGGTGTGGAGGGA	273		
QY 409	gaacgcttcagaagggccaactgtctacacaccacagtgatgtgtctgtcgggagctgct	468		
Db 272	AATCGTTCCTCCAGGAGCTGTGCTCGGCTGCCACGCTGCCAAAGTCAAGAAATAAGCT	213		
QY 469	gccaaagtattacagg 484			

RESULT	9				
LOCUS	AC005612				
DEFINITION	AC005612 Homo sapiens chromosome 21, p1 clone LBL#8 (LBNL H8), complete sequence.				
ACCESSION	AC005612				
NID	93540153				
KEYWORDS	HTG.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 60904)				
TITLE	Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kader,K., Miguel,T., Miller,C., Piltuck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.				
JOURNAL	Sequencing of human chromosome 21				
AUTHORS	Unpublished				
TITLE	2 (bases 1 to 60904)				
REFERENCE	Ricke,D.O.				
AUTHORS	Large Scale Sequence Analysis and Annotation with the Sequences Comparison Analysis (SCAN) System				
TITLE	Unpublished				
JOURNAL	3 (bases 1 to 60904)				
AUTHORS	Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kader,K., Miguel,T., Piltuck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.				
TITLE	Direct Submision				
JOURNAL	Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.				
COMMENT	Sequence submitted by:				
FEATURES	DOE Joint Genome Institute.				
SOURCE	Location/Qualifiers				
	1. .60904				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/map="21"				
	/clone="P1 LBL#8"				
	/chromosome="21"				
	/note="LBNL H8"				
	1. .675				
	/rpt_family="MER25"				
	423. .565				
	/rpt_family="L1"				
	815. .4661				
	/rpt_family="L1"				
	4917. 4940				
	/note="(ACT)8"				
	/rpt_type=tandem				
	/rpt_unit=ACT				
	complement(5525. .5890)				
	/rpt_family="THEL"				
	complement(6688. .6619)				
	/rpt_family="Alu"				
	7441. .7741				
	/rpt_family="Alu"				
	7724. .7743				
	/note="(A)20"				
	/rpt_type=tandem				
	/rpt_unit=A				
	complement(8063. .8123)				
	/rpt_family="MER5"				
	complement(8811. .9778)				
	/rpt_family="LTR5"				
	complement(9970. .10409)				
	/rpt_family="MER41"				
	complement(11266. .11344)				
	/rpt_family="LTR12"				

```
repeat_region 12169..12204
/note="(GT)18"
/rpt_type=tandem
/rpt_unit=GT
repeat_region complement(13074..13520)
/rpt_family="MER7"
repeat_region complement(13974..14029)
/rpt_family="MERS"
repeat_region 15399..15473
/rpt_family="L1"
misc_feature 15697..15808
/note="GRAIL 2 excellent exon, frame 0"
complement(15858..16136)
/rpt_family="Alu"
misc_feature 16565..16698
/note="GRAIL 2 excellent exon, frame 1"
17631..17863
/rpt_family="Alu"
repeat_region complement(18991..19253)
/rpt_family="Alu"
repeat_region complement(19424..19715)
/rpt_family="Alu"
20241..20286
/note="GRAIL 2 excellent exon, frame 0"
21081..21263
/rpt_family="MERS"
21861..22013
/rpt_family="MER20"
repeat_region complement(23391..23728)
/rpt_family="THE1"
23779..23994
/note="GRAIL 2 excellent exon, frame 0"
complement(24012..24087)
/rpt_family="MLT1"
24957..25094
/note="GRAIL 2 excellent exon, frame 1"
25319..25585
/rpt_family="Alu"
27635..27825
/note="GRAIL 2 excellent exon, frame 0"
28141..28430
/rpt_family="Alu"
28834..28966
/note="GRAIL 2 excellent exon, frame 0"
29686..29929
/note="GRAIL 2 excellent exon, frame 2"
29977..30000
/note="(GT)12"
/rpt_type=tandem
/rpt_unit=GT
31701..31844
/note="GRAIL 2 excellent exon, frame 2"
complement(31932..32057)
/rpt_family="MERS"
32297..32320
/note="(T)24"
/rpt_type=tandem
/rpt_unit=T
complement(32303..32508)
/rpt_family="Alu"
33424..33584
/note="GRAIL 2 excellent exon, frame 0"
33976..34117
/note="GRAIL 2 excellent exon, frame 0"
complement(34362..34558)
/rpt_family="trigger1"
complement(34577..34696)
/rpt_family="trigger1"
complement(35034..35111)
/rpt_family="MLT1"
complement(35713..36022)
/rpt_family="Alu"
36501..36803
```

```
repeat_region /rpt_family="Alu"
36909..36950
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
misc_feature 37127..37266
/note="GRAIL 2 excellent exon, frame 2"
39543..39822
/rpt_family="Alu"
misc_feature 40592..40840
/note="GRAIL 2 excellent exon, frame 1"
complement(40971..41337)
/rpt_family="THE1"
repeat_region complement(41617..41717)
/rpt_family="MER4"
repeat_region complement(42881..43215)
/rpt_family="Alu"
43942..44219
/rpt_family="Alu"
repeat_region complement(44724..45031)
/rpt_family="Alu"
repeat_region complement(46007..46178)
/rpt_family="Alu"
46500..46725
/note="GRAIL 2 excellent exon, frame 2"
47500..47781
/rpt_family="Alu"
misc_feature complement(49242..49309)
/note="GRAIL 2 excellent exon, frame 2"
49935..50077
/rpt_family="THE1"
50080..51967
/rpt_family="MSTAR"
repeat_region complement(50244..50543)
/rpt_family="Alu"
misc_feature complement(51595..51645)
/note="GRAIL 2 excellent exon, frame 1"
51968..52329
/rpt_family="THE1"
repeat_region 55080..55378
/rpt_family="Alu"
misc_feature complement(56369..56511)
/note="GRAIL 2 excellent exon, frame 0"
complement(58621..58678)
/note="GRAIL 2 excellent exon, frame 0"
59339..59470
/rpt_family="MER20"
misc_feature complement(59793..59910)
/note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN

Query Match 5.8%; Score 36; DB 42; Length 60904;
Best Local Similarity 49.0%; Pred. No. 3;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 289 gtccaaggaatgctgctcgtcgtatgaggagagacagcagtgctgacctgcaagccgca 348
Db 16307 GTACATGATTAAAGGAGGTATAGGTAAGACCCACAGTGAGGCTCTGGAGCCACGC 16366

QY 349 ccggttccaaggaagactggtgttcacaagaagtgaagccatgctgcgactgtg 408
Db 16367 CCACGTGACAGCCACTCCAGGAGAGTCCAAATCCCGCTAGTGTGGGTGTGAGGGA 16426

QY 409 gaacgcgttccaagggcgaactcgtcacacacagtgatgctgtctgcgggactgcct 468
Db 16427 AATCTGTTCTCCACAGGACCTCGTCTGCGCTGCCACAGCTCCAAAGTCAAGAAATACCT 16486

QY 469 gccaggaattacccg 484
Db 16487 TTCAAAATCTCAGTG 16502
```

NID	92673898
KEYWORDS	Htg.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 140977) Evans,G.A., Athanasiou,M., Basil,M., Bradbury,P., Brignac,S., Bunester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
REFERENCE	template
AUTHORS	Unpublished 2 (bases 1 to 140977) Evans,G.A., Bradbury,P., Brignac,S., Bunester,R., Burbee,D., Davis,J., Davies,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
TITLE	Direct Submission
JOURNAL	Submitted (11-SEP-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA 3 (bases 1 to 140977) Evans,G.A., Athanasiou,M., Basil,M., Bradbury,P., Brignac,S., Bunester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
TITLE	Direct Submission
JOURNAL	Submitted (10-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA On Dec 10, 1997 this sequence version replaced gi:2392758.
COMMENT	
FEATURES	Location/Qualifiers 1..140977 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="PDJ1075120" /chromosome="11"
SOURCE	
BASE COUNT	34355 a 37819 c 37235 g 31568 t
ORIGIN	
Query Match	5.7%; Score 35.6; DB 11; Length 140977;
Best Local Similarity	50.0%; Pred. No. 4.3;
Matches	89; Conservative % 0; Mismatches 89; Indels 0; Gaps 0;
Dy	220 gcagcagaatccaagtgcgatctgtaaacatgtgtctctcgcaaacagtcggacctgg 279
Dy	55124 GGAGAGCAATTCACGACCACCGGGCAGAGAGCCGCCGTTAGCACCTCTGGAGGG 55183
Dy	280 catgaagtgtccaaagatgtgcttgcgttcgtatcggaggagtagcacagtgtgcccctg 339
Dy	55184 CCGGAGTGTCTCCGATGATGATTACCTGGAAACATTTTACACCTCTTCAGGCCAC 55243
Dy	340 caggcgcgcagcgttccaagagactgggtttccagaagtgtatgccatgttcggacc 397
Dy	55244 COTGCCCAAGAGGCCCAAGGAACCTGCCTACTGGGGCCGAGCGTGCCCCC 55301
RESULT	12
LOCUS	AC002536 140977 bp DNA PRI 10-DEC-1997
DEFINITION	Human Chromosome 11 pac PDJ1075120, complete sequence.
ACCESSION	AC002536
NID	92673898
KEYWORDS	Htg.

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eumariotae: Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 140977) Evans, G.A., Athanaslou, M., Basil, M., Bradbury, P., Bignac, S., Bumester, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Nareyanaswamy, U., Newton, J., O'Brien, K., Patel, P., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
TITLE	template
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 140977) Evans, G.A., Bradbury, P., Bignac, S., Bumester, R., Burbee, D., Davle, J., Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L., Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Megarity, C., Nareyanaswamy, U., Newton, J., O'Brien, K., Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
AUTHORS	Direct Submission
REFERENCE	Submitted (11-SEP-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
AUTHORS	3 (bases 1 to 140977) Evans, G.A., Athanaslou, M., Basil, M., Bradbury, P., Bignac, S., Bumester, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Nareyanaswamy, U., Newton, J., O'Brien, K., Patel, P., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
TITLE	Direct Submission
JOURNAL	Submitted (10-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT	On Dec 10, 1997 this sequence version replaced gi:2392758.
FEATURES	Location/Qualifiers
SOURCE	1..140977 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="pD1075f20" /chromosome="11"
BASE COUNT	34355 a 37819 c 37235 g 31568 t
ORIGIN	
Query Match	5.7%; Score 35.6; DB 41; Length 140977;
Best Local Similarity	50.0%; Pred. No. 4.3;
Matches	89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
OY	220 gcagcaggaatcaagatgcgatctcggaacatgcttcctctgcaaacagtggagactg 279
Db	55124 GGAGGAGCATTTCCAGCACCGGAGGGGCGGCGAGAGCGGCTCTTACACTCTTGGAGGG 55183
OY	280 catggaagtgtccaagaaatgtgctctgcatggaatggaagaaatgacagtgfgccctg 339
Db	55184 CCGGAGTGTGTCCTCCGATGATGATGATACCTCGGAACATTTTCACCCTTTCAAGGCCAC 55243
OY	340 cagcgcgcacccggtccaagaaactggaagtggtttccagaagtgtlaagcactgtgcgac 397
Db	55244 CTTGCCCGCAGAGGCCACGAGAAACCTGCTTACTAGGGGGCGCAGCGCTGCCCCAC 55301
RESULT	13
LOCUS	AC005829 170030 bp DNA PRI 15-NOV-1998
DEFINITION	Homo sapiens chromosome 17, clone hRPK.259_G_18, complete sequence
ACCESSION	AC005829
NID	G3873300
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	JOURNAL	TITLE
Eukaryota: Metazoa; Chordata; Vertebrata: Mammalia: Eutheria:		
Primates: Catarrhini: Homnidae; Homo.		
1 (bases 1 to 170030)	Birren,B., Linton,L., Nusbaum,C. and Lander,E.	Submitted (17-Oct-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
Homo sapiens chromosome 17, clone hRRK.259_G_18 unpublished	(bases 1 to 170030)	
2 (bases 1 to 170030)	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckertly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Doneilan,L., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Hearford,A., Herena,L., Horton,L., Howland,J.C., Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczyk,J., Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K., Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychalack,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stillwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	
Direct Submission		
Submitted (31-Nov-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
3 (bases 1 to 170030)	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckertly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Doneilan,L., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Hearford,A., Herena,L., Horton,L., Howland,J.C., Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczyk,J., Macdonald,P., Margulis,N., McEwan,P., McGurk,A., McKernan,K., Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychalack,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stillwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	
Direct Submission		
Submitted (15-Nov-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
On Nov 15, 1998 this sequence version replaced g1:386742.		
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)		
http://ftp.genome.washington.edu/RM/RepeatMasker.html		
FEATURES		
source	Location/Qualifiers	
repeat_region	1..170030	/organism="Homo sapiens"
repeat_region	476..786	/db_xref="taxon:9606"
repeat_region	1750..1915	/clone="hRRK.259_G_18"
repeat_region	2..271	/clone_lib="RPCT-II human BAC library"
repeat_region	476..786	/map="17"
repeat_region	1750..1915	/chromosome="17"
repeat_region	2..271	
repeat_region	476..786	/rpt_family="LJME"
repeat_region	1750..1915	/rpt_family="Aluub"
repeat_region	2..271	/rpt_family="12"
repeat_region	476..786	complement(2143..2448)
repeat_region	1750..1915	/rpt_family="Aluy"
repeat_region	2..271	complement(2600..2810)
repeat_region	476..786	/rpt_family="MIR"
repeat_region	1750..1915	complement(5987..6018)
repeat_region	2..271	/rpt_family="(CCCTA)n"
repeat_region	476..786	6704..6987

```
repeat_region /rpt_family="Alusq"
6988. .7007 /rpt_family="(TAAA)n"
repeat_region 7141. .7173 /rpt_family="AT_rich"
repeat_region complement(7360. .7633)
/rpt_family="Tiggers(Golem)"
repeat_region 7634. .7942 /rpt_family="Alusx"
complement(7943. .8051)
repeat_region /rpt_family="Tiggers(Golem)"
complement(8528. .8838)
repeat_region /rpt_family="Alusx"
complement(8901. .9240)
repeat_region /rpt_family="Alub"
10142. .10422 /rpt_family="Alusq"
repeat_region /rpt_family="Alub"
10473. .10749 /rpt_family="Alub"
repeat_region complement(10750. .10769)
/rpt_family="(CAAAA)n"
repeat_region complement(10771. .11044)
/rpt_family="Alub"
repeat_region 11249. .11366 /rpt_family="Alusq/X"
repeat_region 11371. .11391 /rpt_family="AT_rich"
repeat_region 11417. .11703 /rpt_family="Alub"
repeat_region 11888. .12188 /rpt_family="Alusq"
repeat_region 12189. .12263 /rpt_family="MER41B"
repeat_region 12275. .12290 /rpt_family="Alub"
repeat_region 12291. .12582 /rpt_family="Alusq"
repeat_region 12583. .12731 /rpt_family="Alub"
repeat_region 13548. .13659 /rpt_family="Alub"
repeat_region 15721. .16004 /rpt_family="Alub"
repeat_region 16307. .17274 /rpt_family="Alub"
repeat_region 17651. .17753 /rpt_family="LTR5"
repeat_region /rpt_family="L1MC/D"
complement(17754. .18051)
repeat_region /rpt_family="Alusq"
18052. .18063 /rpt_family="L1MC/D"
repeat_region 18167. .18481 /rpt_family="Alusx"
repeat_region 18486. .18796 /rpt_family="Alub"
repeat_region 18846. .18873 /rpt_family="Alub"
repeat_region /rpt_family="AT_rich"
18945. .19239 /rpt_family="Alusx"
repeat_region 19243. .19509 /rpt_family="Alub"
repeat_region complement(19558. .19700)
/rpt_family="L2"
repeat_region 19709. .19968 /rpt_family="HAL1"
repeat_region 20013. .20318 /rpt_family="Alusq"
repeat_region 20343. .20369 /rpt_family="AT_rich"
complement(20370. .20495)
repeat_region /rpt_family="FLAM_C"
complement(20502. .20797)
/rpt_family="Alusq"

repeat_region 20798. .20815 /rpt_family="AT_rich"
repeat_region 20866. .21093 /rpt_family="HAL1"
repeat_region 21094. .21398 /rpt_family="Alub"
repeat_region 21405. .21715 /rpt_family="Alusx"
repeat_region 21716. .21785 /rpt_family="HAL1"
repeat_region 21791. .22093 /rpt_family="Alusq"
repeat_region 22187. .22465 /rpt_family="Alusq"
repeat_region 22466. .22521 /rpt_family="(CAAAA)n"
repeat_region 22643. .22938 /rpt_family="Alusq"
repeat_region 22947. .23251 /rpt_family="Alusq"
repeat_region 23784. .24097 /rpt_family="Alub"
repeat_region 24698. .24993 /rpt_family="Alusq"
repeat_region 25122. .25147 /rpt_family="(CA)n"
repeat_region 25698. .26013 /rpt_family="Alusq"
repeat_region 26078. .26104 /rpt_family="AT_rich"
repeat_region complement(27061. .27091)
/rpt_family="(CA)n"
repeat_region 28173. .28481 /rpt_family="Alusq"
repeat_region 28829. .29117 /rpt_family="Alusq"
repeat_region 29407. .29727 /rpt_family="Alusq"
repeat_region /rpt_family="Alusq"
complement(2974. .30261)
/rpt_family="Alusq"
31297. .31320 /rpt_family="AT_rich"
repeat_region complement(31353. .31636)
/rpt_family="Alusq"
31885. .31907 /rpt_family="AT_rich"
repeat_region 32571. .32718 /rpt_family="Alusq"
repeat_region 32824. .32994 /rpt_family="Alusq"
repeat_region 33705. .33731 /rpt_family="GC_rich"
complement(33856. .33971)
/rpt_family="(CCGCG)n"
repeat_region 34016. .34043 /rpt_family="GC_rich"
complement(34842. .35072)
/rpt_family="Alusq"
35111. .35419 /rpt_family="Alusq"
repeat_region 35420. .35517 /rpt_family="purine-rich"
repeat_region 35518. .35638 /rpt_family="(GAA)n"
complement(36572. .36603)
/rpt_family="L2"
repeat_region 36626. .36741 /rpt_family="FLAM_C"
repeat_region 36779. .36821 /rpt_family="MER7"
complement(36849. .36877)
/rpt_family="(TAAA)n"
repeat_region 36892. .37008
```

repeat_region	/rpt_family="AluSg/x"	repeat_region	46451..46500
repeat_region	complement(37014..37315)	repeat_region	/rpt_family="(TAA)n"
repeat_region	/rpt_family="AluSp".37618)	repeat_region	complement(46504..46697)
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="HERVL"
repeat_region	complement(37643..37671)	repeat_region	complement(46712..46811)
repeat_region	/rpt_family="(CAA)n"	repeat_region	/rpt_family="MERVL"
repeat_region	complement(37672..37955)	repeat_region	complement(47100..48514)
repeat_region	/rpt_family="AluUb"	repeat_region	/rpt_family="SVA"
repeat_region	complement(38073..38228)	repeat_region	48524..48673
repeat_region	/rpt_family="AluSp/q"	repeat_region	/rpt_family="AluSg/x"
repeat_region	38377..38500	repeat_region	complement(49245..49330)
repeat_region	/rpt_family="FLAM_C"	repeat_region	/rpt_family="MER20"
repeat_region	38501..38528	repeat_region	49331..49629
repeat_region	/rpt_family="(GAA)n"	repeat_region	/rpt_family="AluSp"
repeat_region	38529..38630	repeat_region	complement(49630..49711)
repeat_region	/rpt_family="FLAM_C"	repeat_region	/rpt_family="MER20"
repeat_region	38639..38926	repeat_region	complement(49773..49870)
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="AluO/FRAM"
repeat_region	39001..39036	repeat_region	49952..50063
repeat_region	/rpt_family="AT-rich"	repeat_region	/rpt_family="purine-rich"
repeat_region	complement(39115..39152)	repeat_region	complement(50064..50135)
repeat_region	/rpt_family="(GAA)n"	repeat_region	/rpt_family="AluJ/FLAM"
repeat_region	39176..39471	repeat_region	complement(50136..50563)
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="MUT2A"
repeat_region	39486..39644	repeat_region	50658..50789
repeat_region	/rpt_family="MUT1F"	repeat_region	/rpt_family="AluSg/x"
repeat_region	complement(39673..39963)	repeat_region	50792..51074
repeat_region	/rpt_family="L1PB1"	repeat_region	/rpt_family="AluSg"
repeat_region	39970..40272	repeat_region	complement(51313..51533)
repeat_region	/rpt_family="AluSp"	repeat_region	/rpt_family="L2"
repeat_region	41835..41868	repeat_region	complement(51552..51673)
repeat_region	/rpt_family="POLY_A"	repeat_region	/rpt_family="AluSg/x"
repeat_region	41871..42183	repeat_region	51674..51713
repeat_region	/rpt_family="AluY"	repeat_region	/rpt_family="LIME1"
repeat_region	complement(42184..42379)	repeat_region	51714..52008
repeat_region	/rpt_family="L1PB1"	repeat_region	/rpt_family="MER2"
repeat_region	42380..42671	repeat_region	52009..52068
repeat_region	/rpt_family="AluSc"	repeat_region	/rpt_family="LIME1"
repeat_region	complement(42672..42691)	repeat_region	52069..52361
repeat_region	/rpt_family="L1PB1"	repeat_region	/rpt_family="AluSx"
repeat_region	42754..43049	repeat_region	52362..52652
repeat_region	/rpt_family="MUT1F"	repeat_region	/rpt_family="LIME1"
repeat_region	complement(43054..43356)	repeat_region	complement(53828..53857)
repeat_region	/rpt_family="AluO"	repeat_region	/rpt_family="(CAA)n"
repeat_region	43363..43393	repeat_region	complement(53858..54129)
repeat_region	/rpt_family="AT-rich"	repeat_region	/rpt_family="AluSx"
repeat_region	complement(43400..43414)	repeat_region	54335..54464
repeat_region	/rpt_family="MTR".43723)	repeat_region	/rpt_family="L2"
repeat_region	complement(43415..43723)	repeat_region	complement(54679..54876)
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="MTR"
repeat_region	complement(43724..43884)	repeat_region	complement(54955..55100)
repeat_region	/rpt_family="MTR"	repeat_region	/rpt_family="polypurine"
repeat_region	complement(44278..44464)	repeat_region	57135..57172
repeat_region	/rpt_family="MUT1A1"	repeat_region	/rpt_family="GC-rich"
repeat_region	complement(44465..44774)	repeat_region	57259..57279
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="GC-rich"
repeat_region	complement(44775..44909)	repeat_region	complement(57293..57683)
repeat_region	/rpt_family="MUT1A1"	repeat_region	/rpt_family="purine-rich"
repeat_region	44914..45213	repeat_region	complement(57790..57927)
repeat_region	/rpt_family="AluY"	repeat_region	/rpt_family="L2"
repeat_region	45223..45527	repeat_region	complement(58042..58151)
repeat_region	/rpt_family="AluSp"	repeat_region	/rpt_family="MER21B"
repeat_region	45620..45752	repeat_region	58233..58473
repeat_region	/rpt_family="AluUb"	repeat_region	/rpt_family="AluUb"
repeat_region	45753..46028	repeat_region	58474..58752
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="AluSg1"
repeat_region	46031..46055	repeat_region	complement(58929..59240)
repeat_region	/rpt_family="(CAAA)n"	repeat_region	/rpt_family="AluO"
repeat_region	46056..46342	repeat_region	complement(59340..59645)
repeat_region	/rpt_family="AluSg"	repeat_region	/rpt_family="AluSp"
repeat_region	46343..46421	repeat_region	complement(59652..59791)
repeat_region	/rpt_family="AluUb"	repeat_region	/rpt_family="FLAM_C"
repeat_region		repeat_region	complement(59794..59841)

```

repeat_region /rpt_family="(TAAA)n"
complement(59842..60119)
repeat_region /rpt_family="Alusx"
60120..60145
/rpt_family="AT-rich"
repeat_region complement(60269..60398)
/rpt_family="MLT1A2"
repeat_region complement(60399..60502)
/rpt_family="MER6B"
complement(60503..60788)
/rpt_family="Alusx"
repeat_region complement(60789..61024)
/rpt_family="MER6B"
61025..61312
/rpt_family="Alusg"
complement(61313..61485)
/rpt_family="MER6B"
repeat_region complement(61486..61639)
/rpt_family="MLT1A2"
complement(61647..61815)
/rpt_family="Alusx"
repeat_region complement(61816..62123)
/rpt_family="AluB8"
repeat_region complement(62124..62275)
/rpt_family="Alusx"
62276..62545
/rpt_family="L1M4"
62571..62662
/rpt_family="Alusg"
complement(62663..63313)
/rpt_family="SVA"
63314..63369
/rpt_family="Alusg"
63370..63601
/rpt_family="Alusg"
repeat_region 63604..63708
/rpt_family="L1M4"
complement(63785..64085)
/rpt_family="AluY"
64101..64161
/rpt_family="L1M4"
complement(64162..64333)
/rpt_family="FAM"
64334..64837
/rpt_family="L1M4"
complement(64838..64988)
/rpt_family="AluB"
repeat_region complement(64989..65282)
/rpt_family="AluSp"
complement(65283..65430)
/rpt_family="AluB"
65431..65646
/rpt_family="L1M4"
complement(65647..65815)
/rpt_family="AlusC"
complement(65816..66116)
/rpt_family="AluSp"
complement(66117..66161)
/rpt_family="AlusC"
66162..66229
/rpt_family="L1M4"
66232..66302
/rpt_family="L1ME"
66325..66368
/rpt_family="Alusg/x"
66369..66671
/rpt_family="Alusx"
66672..66852
/rpt_family="Alusg/x"
66909..66915
/note="Single-stranded coverage."
66912..67000
/rpt_family="HAL1"

repeat_region 67001..67305
/rpt_family="Alusx"
67306..67363
/rpt_family="HAL1"
repeat_region complement(67372..67530)
/rpt_family="FAM"
67539..67688
/rpt_family="AluB/FLAM"
complement(67689..68034)
/rpt_family="Alusx"
complement(68036..68328)
/rpt_family="Alusg"
complement(68427..68721)
/rpt_family="Alusg"
68750..69186
/rpt_family="L2"
69247..69346
/rpt_family="MER5A"
complement(69347..69551)
/rpt_family="Alusg"
69552..69859
/rpt_family="AluY"
complement(69860..69985)
/rpt_family="Alusg"
69986..70026
/rpt_family="MER5A"
70172..70287
/rpt_family="L2"
complement(70291..70564)
/rpt_family="AluB"
70566..70626
/rpt_family="Alu"
70690..70771
/rpt_family="HAL1"
70779..70807
/rpt_family="Alu"
71156..71189
/rpt_family="(TA)n"
complement(71249..71356)
/rpt_family="MIR"
complement(71550..71586)
/rpt_family="tRNA-Tyr-TAC"
complement(71616..71906)
/rpt_family="AluB"
72074..72106
/rpt_family="L2"
72174..72205
/rpt_family="MER4D"
72212..72233
/rpt_family="AT-rich"
complement(72310..72609)
/rpt_family="Alusg"
complement(72863..72880)
/rpt_family="AluB"
complement(72881..72917)
/rpt_family="(CA)n"
complement(72918..73202)
/rpt_family="AluB"
73232..73537
/rpt_family="AluY"
complement(73542..73635)
/rpt_family="MIR"
complement(73717..73994)
/rpt_family="Alusg"
complement(74083..74393)
/rpt_family="Alusx"
74529..74717
/rpt_family="L2"
75055..75100
/rpt_family="AT-rich"
complement(75524..75896)
/rpt_family="L1M3e"
76331..76355

```

repeat_region /rpt_family="AT_rich"
78004. .78036
repeat_region /rpt_family="AT_rich"
79901. .80186
repeat_region /rpt_family="Alusx"
80340. .80372
repeat_region /rpt_family="AT_rich"
80373. .80668
repeat_region /rpt_family="AluYb8"
complement(80759. .81066)
repeat_region /rpt_family="AluY"
81223. .81533
repeat_region /rpt_family="Alusg"
complement(81615.81909)
repeat_region /rpt_family="Alusc"
81911. .82058
repeat_region /rpt_family="MIR"
complement(82092. .82382)
repeat_region /rpt_family="Alusx"
82445. .82612
repeat_region /rpt_family="MIR"
complement(82641. .82805)
repeat_region /rpt_family="MLT1B"
complement(82848. .83145)
repeat_region /rpt_family="Alusx"
complement(83146. .83281)
repeat_region /rpt_family="AluYb"
complement(83282. .83404)
repeat_region /rpt_family="MLT1B"
complement(83405. .83694)
repeat_region /rpt_family="Alusx"
complement(83695. .83756)
repeat_region /rpt_family="MLT1B"
complement(83757. .84065)
repeat_region /rpt_family="AluYb"
complement(84066. .84114)
repeat_region /rpt_family="MLT1B"
complement(84272. .84557)
repeat_region /rpt_family="Alusg"
84563. .84818
repeat_region /rpt_family="L2"
complement(84819. .85129)
repeat_region /rpt_family="Alusx"
85130. .85480
repeat_region /rpt_family="L2"
85717. .85816
repeat_region /rpt_family="HY3"
complement(86213. .86516)
repeat_region /rpt_family="Alusc"
complement(86959. .87253)
repeat_region /rpt_family="Alusx"
87281. .87573
repeat_region /rpt_family="AluYb8"
87574. .87599
repeat_region /rpt_family="(CAA)n"
87600. .87612
repeat_region /rpt_family="AluYb8"
87891. .88354
repeat_region /rpt_family="L1MC/D"
88370. .88663
repeat_region /rpt_family="Alusg"
complement(89376. .89597)
repeat_region /rpt_family="MER30"
complement(90386. .90696)
repeat_region /rpt_family="Alusg"
91095. .91396
repeat_region /rpt_family="Alusx"
complement(91689. .91997)
repeat_region /rpt_family="Alusg"
complement(92020. .92131)
repeat_region /rpt_family="Alusg/x"
complement(92191.92249)
repeat_region /rpt_family="MLT1A2"

repeat_region complement(92251. .92558)
repeat_region /rpt_family="AluYb"
complement(92745. .93059)
repeat_region /rpt_family="Alusg"
complement(93060. .93192)
repeat_region /rpt_family="AluYb"
complement(93193. .93293)
repeat_region /rpt_family="L1M4"
complement(93294. .93353)
repeat_region /rpt_family="AluYb"
complement(93354. .93658)
repeat_region /rpt_family="Alusg1"
complement(93659. .93867)
repeat_region /rpt_family="AluYb"
complement(93888. .94306)
repeat_region /rpt_family="L1M4"
complement(94312. .94443)
repeat_region /rpt_family="FLAM_A"
complement(94445. .94777)
repeat_region /rpt_family="L1M4"
complement(94778. .95092)
repeat_region /rpt_family="Alusg"
complement(95093. .95189)
repeat_region /rpt_family="L1M4"
complement(95190. .95488)
repeat_region /rpt_family="Alusx"
complement(95489. .96317)
repeat_region /rpt_family="L1M4"
complement(96413. .96713)
repeat_region /rpt_family="Alusx"
complement(96714. .96992)
repeat_region /rpt_family="AluYb"
complement(97632. .97938)
repeat_region /rpt_family="Alusg"
complement(98515. .98684)
repeat_region /rpt_family="MER46B"
complement(98685. .98866)
repeat_region /rpt_family="AluYb"
complement(98867. .99159)
repeat_region /rpt_family="AluYb"
complement(99160. .99298)
repeat_region /rpt_family="AluYb"
complement(99299. .99480)
repeat_region /rpt_family="MER46B"
99677. .99978
repeat_region /rpt_family="Alusx"
100115. .100142
repeat_region /rpt_family="AT_rich"
complement(100143. .100427)
repeat_region /rpt_family="AluY"
100428. .100445
repeat_region /rpt_family="AT_rich"
complement(100832. .100849)
repeat_region /rpt_family="(CA)n"
complement(102147. .102436)
repeat_region /rpt_family="Alusx"
complement(102439. .102465)
repeat_region /rpt_family="Poly_A"
complement(102608. .102751)
repeat_region /rpt_family="MIR"
complement(102756. .102859)
repeat_region /rpt_family="L2"
complement(103236. .103527)
repeat_region /rpt_family="Alusx"
complement(103725. .103756)
repeat_region /rpt_family="polypurine"
complement(103757. .104063)
repeat_region /rpt_family="AluYb"
complement(104625. .104928)
repeat_region /rpt_family="AluYb"
105057. .105317
repeat_region /rpt_family="AluYb"
105712. .106734

```
repeat_region /rpt_family="GC_rich"
107291. .107362
/rpt_family="GC_rich"
repeat_region complement(107433. .107457)
/rpt_family="GCAG)n"
107595. .107618
/rpt_family="GC_rich"
repeat_region 107733. .107762
/rpt_family="GC_rich"
repeat_region 107928. .108001
/rpt_family="GC_rich"
repeat_region complement(108073. .108129)
/rpt_family="purine-rich"
repeat_region complement(110962. .111287)
/rpt_family="AluSx"
repeat_region complement(111508. .111578)
/rpt_family="L1M3e"
repeat_region complement(111639. .111804)
/rpt_family="FAM"
repeat_region complement(111827. .112139)
/rpt_family="AluSx"
repeat_region 112907. .112928
/rpt_family="AT_rich"
repeat_region 113857. .114160
/rpt_family="AluSx"
repeat_region complement(114487. .114794)
/rpt_family="AluYb8"
repeat_region complement(115026. .115147)
/rpt_family="FRAM_C"
repeat_region complement(115154. .115233)
/rpt_family="(CA)n"
repeat_region complement(115676. .115761)
/rpt_family="L2"
repeat_region 115762. .116050
/rpt_family="AluSg"
repeat_region complement(116051. .116264)
/rpt_family="L2"
repeat_region 116638. .116700
/rpt_family="L2"
repeat_region complement(116960. .117219)
/rpt_family="AluSx"
repeat_region complement(117220. .117335)
/rpt_family="G-rich"
repeat_region complement(117336. .117590)
/rpt_family="AluYb"
repeat_region complement(117632. .117741)
/rpt_family="FRAM_A"
repeat_region 117742. .117809
/rpt_family="L2"
repeat_region 118052. .118266
/rpt_family="MIR"
repeat_region complement(118930. .119080)
/rpt_family="L2"
repeat_region 119157. .119454
/rpt_family="AluSg"
repeat_region complement(119486. .119609)
/rpt_family="L2"
repeat_region 119798. .119961
/rpt_family="L2"
repeat_region 119993. .120040
/rpt_family="(TA)n"
repeat_region complement(120042. .120170)
/rpt_family="FRAM_C"
repeat_region complement(120180. .120356)
/rpt_family="AluSg/x"
repeat_region complement(120360. .120674)
/rpt_family="AluSg"
repeat_region complement(120809. .121101)
/rpt_family="AluSg"
repeat_region 121680. .121893
/rpt_family="MER20"
repeat_region 122812. .123110
/rpt_family="AluYb8"

repeat_region complement(123636. .123802)
/rpt_family="MIR"
repeat_region complement(124180. .124479)
/rpt_family="AluSg"
repeat_region 125035. .125070
/rpt_family="AT_rich"
repeat_region 125239. .125524
/rpt_family="AluSx"
repeat_region 125525. .125577
/rpt_family="(CA)n"
repeat_region complement(125807. .125875)
/rpt_family="Charlie1"
repeat_region 125903. .125924
/rpt_family="AT_rich"
repeat_region 125933. .126199
/rpt_family="AluYb"
repeat_region complement(126217. .126407)
/rpt_family="Charlie1a"
repeat_region 126408. .126681
/rpt_family="AluSc"
repeat_region complement(126682. .127379)
/rpt_family="Charlie1a"
repeat_region complement(127634. .128101)
/rpt_family="L2"
repeat_region complement(130705. .130741)
/rpt_family="(CAAA)n"
repeat_region complement(130746. .130989)
/rpt_family="AluSg"
repeat_region 132238. .132536
/rpt_family="AluSp"
repeat_region complement(133820. .133882)
/rpt_family="A-rich"
repeat_region complement(134361. .134658)
/rpt_family="AluSg"
repeat_region complement(134660. .134959)
/rpt_family="AluSx"
repeat_region 135078. .135375
/rpt_family="AluSx"
repeat_region 135378. .135669
/rpt_family="AluY"
repeat_region complement(135741. .135752)
/rpt_family="AluSg/x"
repeat_region complement(135753. .135779)
/rpt_family="(CAT)n"
repeat_region complement(135780. .135944)
/rpt_family="AluSg/x"
repeat_region 136851. .136881
/rpt_family="MIR"
repeat_region 137363. .137522
/rpt_family="FRAM"
repeat_region complement(137618. .138412)
/rpt_family="L2"
repeat_region 138422. .138444
/rpt_family="AT_rich"
repeat_region complement(139440. .139755)
/rpt_family="AluY"
repeat_region complement(139820. .140012)
/rpt_family="MIR"
repeat_region 140299. .140579
/rpt_family="AluYb"
repeat_region complement(140756. .140818)
/rpt_family="MIR"
repeat_region 142048. .142358
/rpt_family="AluSg"
repeat_region 142359. .142524
/rpt_family="AluSg/x"
repeat_region complement(143972. .144135)
/rpt_family="AluSg/x"
repeat_region 144187. .144263
/rpt_family="FRAM"
repeat_region 144276. .144298
/rpt_family="AT_rich"
repeat_region complement(144535. .144820)
```

[illegible]

Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G., Hagos, B., Heatford, A., Herena, L., Horton, L., Howland, J. C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Macdonald, P., Marquis, N., McEwan, P., Mychaleckyj, J., Nahf, R., Naylor, J., Nilotf, M., Morris, W., Morrow, J., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Testfaye, S., Tichovolsky, N., Torrella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (17-OCT-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02144, USA
REFERENCE 3 (bases 1 to 170030)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, O., Barna, N., Becker, R., Benn, J., Bouwelle, C., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E., Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G., Hagos, B., Heatford, A., Herena, L., Horton, L., Howland, J. C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Nilotf, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Testfaye, S., Tichovolsky, N., Torrella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (15-NOV-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02144, USA
COMMENT On Nov 15, 1998 this sequence version replaced g1.3868742. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Only the first 170030 base pairs of this clone are being submitted. The remainder overlaps accession number AC004139 (WIGR project L291).

FEATURES
SOURCE Location/Qualifiers
1. 170030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hprk 259_G18"
/clone_lib="RPCI-11 human BAC library"
/map="17"
/chromosome="17"
2. 271
/rpt_family="L1ME"
476..786
/rpt_family="AluB"
1750..1915
/rpt_family="L2"
complement(2143..2448)
/rpt_family="AluY"
complement(2600..2810)
/rpt_family="MIR"
complement(5987..6018)
/rpt_family="(CCCTA)n"
6704..6987
/rpt_family="AluSg"
6988..7007
/rpt_family="(TAAA)n"
7141..7173
/rpt_family="AT_rich"
complement(7360..7633)
/rpt_family="rigger3(Golem)"
7634..7942
/rpt_family="AluSx"
complement(7943..8051)
/rpt_family="rigger3(Golem)"

repeat_region complement(8528..8838)
/rpt_family="AluSx"
repeat_region complement(8901..9240)
/rpt_family="AluB"
repeat_region 10142..10422
/rpt_family="AluSg"
repeat_region 10473..10749
/rpt_family="AluB"
repeat_region complement(10750..10769)
/rpt_family="(CAAA)n"
complement(10771..11044)
/rpt_family="AluB"
repeat_region 11249..11366
/rpt_family="AluSg/X"
11371..11391
/rpt_family="AT_rich"
11417..11703
/rpt_family="AluB"
repeat_region 11888..12188
/rpt_family="AluSg"
12189..12263
/rpt_family="MER41B"
12275..12290
/rpt_family="AluB"
12391..12582
/rpt_family="AluSg"
12583..12731
/rpt_family="AluB"
15348..15659
/rpt_family="AluYB8"
15721..16004
/rpt_family="AluO"
16307..17274
/rpt_family="L7R5"
17651..17753
/rpt_family="L1MC/D"
complement(17754..18051)
/rpt_family="AluSc"
18052..18063
/rpt_family="L1MC/D"
18167..18481
/rpt_family="AluSx"
18486..18796
/rpt_family="AluY"
18846..18873
/rpt_family="AT_rich"
18945..19239
/rpt_family="AluSx"
19243..19509
/rpt_family="AluO"
complement(19558..19700)
/rpt_family="L2"
19709..19968
/rpt_family="HAL1"
20013..20318
/rpt_family="AluSg"
20343..20369
/rpt_family="AT_rich"
complement(20370..20495)
/rpt_family="FLAM_C"
complement(20502..20797)
/rpt_family="AluSg"
20798..20815
/rpt_family="AT_rich"
20866..21093
/rpt_family="HAL1"
21094..21398
/rpt_family="AluB"
21405..21715
/rpt_family="AluSx"
21716..21785
/rpt_family="HAL1"
21791..22093
repeat_region


```
repeat_region /rpt_family="Alusg" 22187 . 22465
/rpt_family="Alusg1" 22466 . 22521
repeat_region /rpt_family="(CAAA)n" 22643 . 22938
/rpt_family="Alusx" 22947 . 23251
repeat_region /rpt_family="Alusg" 23784 . 24097
/rpt_family="Alubd" 24698 . 24993
repeat_region /rpt_family="Alusg" 25122 . 25147
/rpt_family="(CA)n" 25698 . 26013
repeat_region /rpt_family="Alusp" 26078 . 26104
/rpt_family="AT_rich" 27061 . 27091
repeat_region /rpt_family="(CAAA)n" 28173 . 28481
/rpt_family="Alusg" 28829 . 29117
repeat_region /rpt_family="Alusg" 29407 . 29727
/rpt_family="Alusp" 29974 . 30261
repeat_region /rpt_family="Alusg" 31297 . 31320
/rpt_family="AT_rich" 31353 . 31366
repeat_region /rpt_family="Alusp" 31885 . 31907
/rpt_family="AT_rich" 32571 . 32718
repeat_region /rpt_family="Alusx" 32824 . 32994
/rpt_family="Alusx" 33705 . 33731
repeat_region /rpt_family="GC_rich" 33856 . 33971
/rpt_family="(CCGC)n" 34016 . 34043
repeat_region /rpt_family="GC_rich" 34842 . 35072
/rpt_family="Alusx" 35111 . 35419
repeat_region /rpt_family="Alusg" 35420 . 35517
/rpt_family="purine-rich" 35518 . 35638
/rpt_family="(GAA)n" 36572 . 36603
repeat_region /rpt_family="I2" 36626 . 36741
/rpt_family="FLAM_C" 36779 . 36821
/rpt_family="MER77" 36892 . 37008
/rpt_family="Alusg/x" 37014 . 37315
repeat_region /rpt_family="Alusp" 37316 . 37618
/rpt_family="Alusx" 37643 . 37671
repeat_region /rpt_family="(CAAA)n" 37672 . 37955
/rpt_family="Alubd" 38073 . 38228
/rpt_family="Alusp/q" 38377 . 38500
repeat_region /rpt_family="FLAM_C" 38501 . 38528
/rpt_family="(GAA)n" 38529 . 38630
repeat_region /rpt_family="FLAM_C" 38639 . 38926
/rpt_family="Alusx" 39001 . 39036
repeat_region /rpt_family="AT_rich" 39115 . 39152
/rpt_family="(GAA)n" 39176 . 39471
repeat_region /rpt_family="Alusx" 39486 . 39644
/rpt_family="MLTIF" 39673 . 39963
repeat_region /rpt_family="LIPB1" 40272
39970 . 40272
repeat_region /rpt_family="Alusp" 41835 . 41868
/rpt_family="POLY_A" 41871 . 42183
/rpt_family="Aluy" 42184 . 42379
repeat_region /rpt_family="LIPB1" 42380 . 42671
/rpt_family="Alusc" 42672 . 42691
repeat_region /rpt_family="LIPB1" 42754 . 43049
/rpt_family="MLTIF" 43054 . 43356
repeat_region /rpt_family="Alubd" 43363 . 43393
/rpt_family="AT_rich" 43400 . 43414
/rpt_family="MIR" 43415 . 43723
/rpt_family="Alusx" 43724 . 43884
/rpt_family="MIR" 44278 . 44464
/rpt_family="MLT1A1" 44465 . 44774
/rpt_family="Alusx" 44775 . 44909
/rpt_family="MLT1A1" 44914 . 45213
/rpt_family="Aluy" 45223 . 45527
/rpt_family="Alusp" 45620 . 45752
repeat_region /rpt_family="Alubd" 45753 . 46028
/rpt_family="Alusx" 46031 . 46055
/rpt_family="(CAAA)n" 46056 . 46342
/rpt_family="Alusg" 46343 . 46421
/rpt_family="Alubd" 46451 . 46500
/rpt_family="TAAAn" 46504 . 46697
/rpt_family="HERVL" 46712 . 46811
/rpt_family="MERVL" 47100 . 48514
/rpt_family="SVA" 48524 . 48673
/rpt_family="Alusg/x" 49245 . 49330
repeat_region complement(49245 . 49330)
```

```

repeat_region      /rpt_family="MER20"
                    49331. .49629
repeat_region      /rpt_family="AluSp"
                    complement(49630. .49711)
repeat_region      /rpt_family="MER20"
                    complement(49773. .49870)
repeat_region      /rpt_family="AluJo/FRAM"
                    49952. .50063
repeat_region      /rpt_family="Purine-rich"
                    complement(50064. .50135)
repeat_region      /rpt_family="AluJo/FLAM"
                    complement(50136. .50563)
repeat_region      /rpt_family="MLT2A"
                    50658. .50789
repeat_region      /rpt_family="AluSg/x"
                    50792. .51074
repeat_region      /rpt_family="AluSg"
                    complement(51313. .51533)
repeat_region      /rpt_family="L2"
                    complement(51552. .51673)
repeat_region      /rpt_family="AluSg/x"
                    51674. .51713
repeat_region      /rpt_family="L1ME1"
                    51714. .52008
repeat_region      /rpt_family="MER2"
                    52009. .52068
repeat_region      /rpt_family="L1ME1"
                    52069. .52361
repeat_region      /rpt_family="AluSx"
                    52362. .52652
repeat_region      /rpt_family="L1ME1"
                    complement(53828. .53857)
repeat_region      /rpt_family="(CAAAA)n"
                    complement(53858. .54123)
repeat_region      /rpt_family="AluSx"
                    54339. .54484
repeat_region      /rpt_family="L2"
                    complement(54679. .54876)
repeat_region      /rpt_family="MIR"
                    complement(54955. .55100)
repeat_region      /rpt_family="polypurine"
                    57139. .57172
repeat_region      /rpt_family="GC-rich"
                    57259. .57279
repeat_region      /rpt_family="GC-rich"
                    complement(57293. .57683)
repeat_region      /rpt_family="Purine-rich"
                    complement(57790. .57927)
repeat_region      /rpt_family="L2"
                    complement(58042. .58151)
repeat_region      /rpt_family="MER21B"
                    58233. .58473
repeat_region      /rpt_family="AluJb"
                    58474. .58752
repeat_region      /rpt_family="AluSg1"
                    complement(58929. .59240)
repeat_region      /rpt_family="AluJo"
                    complement(59340. .59645)
repeat_region      /rpt_family="AluSp"
                    complement(59652. .59791)
repeat_region      /rpt_family="FLAM.C"
                    complement(59794. .59841)
repeat_region      /rpt_family="(TAAA)n"
                    complement(59842. .60119)
repeat_region      /rpt_family="AluSx"
                    60120. .60145
repeat_region      /rpt_family="AT-rich"
                    complement(60269. .60398)
repeat_region      /rpt_family="MLT1A2"
                    complement(60399. .60502)
repeat_region      /rpt_family="MER6B"
                    complement(60503. .60788)
repeat_region      /rpt_family="AluSx"

repeat_region      complement(60789. .61024)
repeat_region      /rpt_family="MER66B"
                    61025. .61312
repeat_region      /rpt_family="AluSg"
                    complement(61313. .61485)
repeat_region      /rpt_family="MER66B"
                    complement(61486. .61639)
repeat_region      /rpt_family="MLT1A2"
                    complement(61647. .61815)
repeat_region      /rpt_family="AluSx"
                    complement(61816. .62123)
repeat_region      /rpt_family="AluYb8"
                    complement(62124. .62275)
repeat_region      /rpt_family="AluSx"
                    62276. .62545
repeat_region      /rpt_family="L1M4"
                    62571. .62662
repeat_region      /rpt_family="AluSg"
                    complement(62663. .63313)
repeat_region      /rpt_family="SVA"
                    63314. .63369
repeat_region      /rpt_family="AluSg"
                    63370. .63601
repeat_region      /rpt_family="AluSg"
                    63604. .63708
repeat_region      /rpt_family="L1M4"
                    complement(63785. .64085)
repeat_region      /rpt_family="AluY"
                    64101. .64161
repeat_region      /rpt_family="L1M4"
                    complement(64162. .64333)
repeat_region      /rpt_family="FRAM"
                    64334. .64837
repeat_region      /rpt_family="L1M4"
                    complement(64838. .64988)
repeat_region      /rpt_family="AluJb"
                    complement(64989. .65282)
repeat_region      /rpt_family="AluSp"
                    complement(65283. .65430)
repeat_region      /rpt_family="AluJb"
                    65431. .65646
repeat_region      /rpt_family="L1M4"
                    complement(65647. .65815)
repeat_region      /rpt_family="AluSc"
                    complement(65816. .66116)
repeat_region      /rpt_family="AluSp"
                    complement(66117. .66161)
repeat_region      /rpt_family="AluSc"
                    66162. .66229
repeat_region      /rpt_family="L1M4"
                    66232. .66302
repeat_region      /rpt_family="L1ME"
                    66325. .66368
repeat_region      /rpt_family="AluSg/x"
                    66369. .66671
repeat_region      /rpt_family="AluSx"
                    66672. .66852
repeat_region      /rpt_family="AluSg/x"
                    66909. .66915
repeat_region      /note="Single-stranded coverage."
                    66912. .67000
repeat_region      /rpt_family="HAL1"
                    67001. .67305
repeat_region      /rpt_family="AluSx"
                    67306. .67363
repeat_region      /rpt_family="HAL1"
                    complement(67372. .67530)
repeat_region      /rpt_family="FRAM"
                    67539. .67688
repeat_region      /rpt_family="AluJo/FLAM"
                    complement(67689. .68034)
repeat_region      /rpt_family="AluSx"
                    complement(68036. .68328)

```

```
repeat_region /rpt_family="Alusq" complement(68427..68721)
repeat_region /rpt_family="Alusq" 68750..69186
repeat_region /rpt_family="L2" 69247..69346
repeat_region /rpt_family="MER5A" complement(69347..69551)
repeat_region /rpt_family="Alusq" 69552..69859
repeat_region /rpt_family="AluY" complement(69860..69985)
repeat_region /rpt_family="Alusq" 69986..70026
repeat_region /rpt_family="MER5A" 70172..70287
repeat_region /rpt_family="L2" complement(70291..70564)
repeat_region /rpt_family="AluDo" 70566..70626
repeat_region /rpt_family="Alu" 70690..70771
repeat_region /rpt_family="MAL1" 70779..70807
repeat_region /rpt_family="Alu" 71156..71189
repeat_region /rpt_family="(TA)n" complement(71249..71356)
repeat_region /rpt_family="MIR" complement(71550..71586)
repeat_region /rpt_family="tRNA-Tyr-TTC" complement(71616..71906)
repeat_region /rpt_family="AluDo" 72074..72106
repeat_region /rpt_family="L2" 72174..72205
repeat_region /rpt_family="MER4D" 72212..72233
repeat_region /rpt_family="AT_rich" complement(72310..72609)
repeat_region /rpt_family="Alusq" complement(72863..72880)
repeat_region /rpt_family="AluDb" complement(72881..72917)
repeat_region /rpt_family="(CAA)n" complement(72918..73202)
repeat_region /rpt_family="AluDb" 73232..73537
repeat_region /rpt_family="AluY" complement(73542..73635)
repeat_region /rpt_family="MIR" complement(73717..73994)
repeat_region /rpt_family="Alusq" complement(74083..74393)
repeat_region /rpt_family="Alusx" 74529..74717
repeat_region /rpt_family="L2" 75055..75100
repeat_region /rpt_family="AT_rich" complement(75524..75896)
repeat_region /rpt_family="L1M3e" 76351..76355
repeat_region /rpt_family="AT_rich" 78004..78036
repeat_region /rpt_family="AT_rich" 79901..80186
repeat_region /rpt_family="Alusx" 80340..80372
repeat_region /rpt_family="AT_rich" 80373..80668
repeat_region /rpt_family="AluYb8" complement(80759..81066)
repeat_region /rpt_family="AluY" 81223..81533
repeat_region /rpt_family="Alusq" complement(81615..81909)
repeat_region /rpt_family="AlusC" 81911..82058
repeat_region /rpt_family="MIR" complement(82092..82382)
repeat_region /rpt_family="Alusx" 82445..82612
repeat_region /rpt_family="MIR" complement(82641..82805)
repeat_region /rpt_family="MT1B" complement(82848..83145)
repeat_region /rpt_family="Alusx" complement(83146..83281)
repeat_region /rpt_family="AluDb" complement(83282..83404)
repeat_region /rpt_family="MT1B" complement(83405..83694)
repeat_region /rpt_family="Alusx" complement(83695..83756)
repeat_region /rpt_family="MT1B" complement(83757..84065)
repeat_region /rpt_family="AluDo" complement(84066..84114)
repeat_region /rpt_family="MT1B" complement(84272..84557)
repeat_region /rpt_family="Alusq" 84563..84818
repeat_region /rpt_family="L2" complement(84819..85129)
repeat_region /rpt_family="Alusx" 85130..85480
repeat_region /rpt_family="L2" 85717..85816
repeat_region /rpt_family="HY3" complement(86213..86516)
repeat_region /rpt_family="AlusC" complement(86559..87253)
repeat_region /rpt_family="Alusx" 87281..87573
repeat_region /rpt_family="AluYb8" 87574..87599
repeat_region /rpt_family="(CAA)n" 87600..87612
repeat_region /rpt_family="AluYb8" 87891..88354
repeat_region /rpt_family="L1MC/D" 88370..88663
repeat_region /rpt_family="Alusq" complement(89376..89597)
repeat_region /rpt_family="MER30" complement(90385..90696)
repeat_region /rpt_family="Alusq" 91095..91396
repeat_region /rpt_family="Alusx" complement(91689..91997)
repeat_region /rpt_family="Alusq" complement(92020..92131)
repeat_region /rpt_family="Alusq/A" complement(92191..92249)
repeat_region /rpt_family="MT1A2" complement(92251..92558)
repeat_region /rpt_family="AluDo" complement(92745..93059)
repeat_region /rpt_family="Alusq" complement(93060..93192)
repeat_region /rpt_family="AluDb" complement(93193..93293)
repeat_region /rpt_family="L1M4" complement(93294..93353)
repeat_region /rpt_family="AluDo" complement(93354..93658)
```

```
repeat_region /rpt_family="AluSg1" complement(93659. .93887)
repeat_region /rpt_family="AluJo" complement(93888. .94306)
repeat_region /rpt_family="L1M4" complement(94312. .94443)
repeat_region /rpt_family="FLAM_A" complement(94445. .94777)
repeat_region /rpt_family="L1M4" complement(94778. .95092)
repeat_region /rpt_family="AluSg" complement(95093. .95189)
repeat_region /rpt_family="L1M4" complement(95190. .95488)
repeat_region /rpt_family="AluSx" complement(95489. .96317)
repeat_region /rpt_family="L1M4" complement(96413. .96713)
repeat_region /rpt_family="AluSx" complement(96714. .96992)
repeat_region /rpt_family="AluJo" complement(97632. .97938)
repeat_region /rpt_family="AluSg" complement(98515. .98684)
repeat_region /rpt_family="MER46B" complement(98685. .98866)
repeat_region /rpt_family="AluJo" complement(98867. .99159)
repeat_region /rpt_family="AluCb" complement(99160. .99298)
repeat_region /rpt_family="AluJo" complement(99299. .99480)
repeat_region /rpt_family="MER46B" 99677. 99878
repeat_region /rpt_family="AluSx" 100115. 100142
repeat_region /rpt_family="AT_rich" complement(100143. .100427)
repeat_region /rpt_family="AluY" 100428. 100445
repeat_region /rpt_family="AT_rich" complement(100822. .100849)
repeat_region /rpt_family="(CA)n" complement(102147. .102436)
repeat_region /rpt_family="AluSx" complement(102439. .102465)
repeat_region /rpt_family="POLY_A" complement(102608. .102751)
repeat_region /rpt_family="MIR" complement(102756. .102859)
repeat_region /rpt_family="L2" complement(103336. .103527)
repeat_region /rpt_family="AluSx" complement(103725. .103756)
repeat_region /rpt_family="polypurine" complement(103757. .104063)
repeat_region /rpt_family="AluJo" complement(104625. .104928)
repeat_region /rpt_family="AluCb" 105057. 105317
repeat_region /rpt_family="AluJo" 106712. 106734
repeat_region /rpt_family="GC_rich" 107291. 107362
repeat_region /rpt_family="GC_rich" complement(107433. .107457)
repeat_region /rpt_family="(CGAG)n" 107595. 107618
repeat_region /rpt_family="GC_rich" 107733. 107762
repeat_region /rpt_family="GC_rich" 107928. 108001
repeat_region /rpt_family="GC_rich"

repeat_region complement(108073. 108129)
repeat_region /rpt_family="purine_rich" complement(110962. .111287)
repeat_region /rpt_family="AluSx" complement(111508. .111578)
repeat_region /rpt_family="L1M3e" complement(111639. .111804)
repeat_region /rpt_family="FAM" complement(111827. .112139)
repeat_region /rpt_family="AluSx" 112907. 112928
repeat_region /rpt_family="AT_rich" 113857. 114160
repeat_region /rpt_family="AluSx" complement(114487. .114794)
repeat_region /rpt_family="AluYb8" complement(115026. .115147)
repeat_region /rpt_family="FLAM_C" complement(115154. .115233)
repeat_region /rpt_family="(CA)n" complement(115676. .115761)
repeat_region /rpt_family="L2" 115762. 116050
repeat_region /rpt_family="AluSg" complement(116051. .116264)
repeat_region /rpt_family="L2" 11638. 116700
repeat_region /rpt_family="L2" complement(116960. .117219)
repeat_region /rpt_family="AluSx" complement(117220. .117335)
repeat_region /rpt_family="GC_rich" complement(117336. .117590)
repeat_region /rpt_family="AluYb" complement(117632. .117741)
repeat_region /rpt_family="FLAM_A" 117742. 117809
repeat_region /rpt_family="L2" 118052. 118266
repeat_region /rpt_family="MIR" complement(118930. .119080)
repeat_region /rpt_family="L2" 119157. 119454
repeat_region /rpt_family="AluSg" complement(119486. .119609)
repeat_region /rpt_family="L2" 119798. 119961
repeat_region /rpt_family="L2" 119993. 120040
repeat_region /rpt_family="(TA)n" complement(120042. .120170)
repeat_region /rpt_family="FLAM_C" complement(120180. .120356)
repeat_region /rpt_family="AluSg/*" complement(120360. .120674)
repeat_region /rpt_family="AluSg" complement(120809. .121101)
repeat_region /rpt_family="AluSg" 121680. 121893
repeat_region /rpt_family="MER20" 122812. 123110
repeat_region /rpt_family="AluYb8" complement(123636. .123802)
repeat_region /rpt_family="MIR" complement(124180. .124479)
repeat_region /rpt_family="AluSg" 125035. 125070
repeat_region /rpt_family="AT_rich" 125239. 125524
repeat_region /rpt_family="AluSx" 125525. 125577
repeat_region /rpt_family="(CA)n" complement(125807. .125875)
```

```
/rpt_family="Charl1el"
repeat_region 125903. .125924
/rpt_family="AT_r1ch"
repeat_region 125933. .126199
/rpt_family="AluUb"
repeat_region complement(126217. .126407)
/rpt_family="Char1elA"
repeat_region 126408. .126681
/rpt_family="AluSc"
repeat_region complement(126682. .127379)
/rpt_family="Char1elA"
repeat_region complement(127634. .128101)
/rpt_family="L2"
repeat_region complement(130705. .130741)
/rpt_family="(CAAA)n"
repeat_region complement(130746. .130989)
/rpt_family="AluSg"
repeat_region 132238. .132536
/rpt_family="AluSp"
repeat_region complement(133820. .133882)
/rpt_family="A_r1ch"
repeat_region complement(134361. .134658)
/rpt_family="AluSg"
repeat_region complement(134660. .134959)
/rpt_family="AluSx"
repeat_region 135078. .135375
/rpt_family="AluSx"
repeat_region 135378. .135669
/rpt_family="AluY"
repeat_region complement(135741. .135752)
/rpt_family="AluSg/x"
repeat_region complement(135753. .135779)
/rpt_family="(CATA)n"
repeat_region complement(135780. .135944)
/rpt_family="AluSg/x"
repeat_region 136851. .136881
/rpt_family="MIR"
repeat_region 137363. .137522
/rpt_family="FRAM"
repeat_region complement(137618. .138412)
/rpt_family="L2"
repeat_region 138422. .138444
/rpt_family="AT_r1ch"
repeat_region complement(139440. .139755)
/rpt_family="AluY"
repeat_region complement(139820. .140012)
/rpt_family="MIR"
repeat_region 140239. .140579
/rpt_family="AluUb"
repeat_region complement(140756. .140818)
/rpt_family="MIR"
repeat_region 142048. .142358
/rpt_family="AluSg"
repeat_region 142359. .142524
/rpt_family="AluSg/x"
repeat_region complement(143972. .144135)
/rpt_family="AluSg/x"
repeat_region 144187. .144263
/rpt_family="FRAM"
repeat_region 144276. .144298
/rpt_family="AT_r1ch"
repeat_region complement(144535. .144820)
/rpt_family="AluSx"
repeat_region complement(144945. .145101)
/rpt_family="FRAM"
repeat_region complement(145210. .145403)
/rpt_family="MER20"
repeat_region 145431. .145732
/rpt_family="AluSp"
repeat_region 145934. .146027
/rpt_family="FRAM_C"
repeat_region 146309. .146538
/rpt_family="AluUb"
repeat_region 148821. .148988
/rpt_family="MER5B"
repeat_region complement(149198. .149678)
/rpt_family="L2"
repeat_region complement(150807. .150838)
/rpt_family="AluYb8"
repeat_region complement(150839. .150864)
/rpt_family="(TAA)n"
repeat_region complement(150865. .151162)
/rpt_family="AluYb8"
repeat_region 151247. .151464
/rpt_family="MER58A"
repeat_region 151588. .151891
/rpt_family="AluY"
repeat_region 151970. .151990
/rpt_family="AT_r1ch"
repeat_region 152003. .152297
/rpt_family="AluSx"
repeat_region complement(152753. .152937)
/rpt_family="MIR"
repeat_region complement(153354. .153291)
/rpt_family="(GA)n"
repeat_region complement(153392. .153595)
/rpt_family="AluUb"
repeat_region 153812. .153834
/rpt_family="AT_r1ch"
repeat_region complement(153843. .153965)
/rpt_family="FRAM_A"
repeat_region complement(154028. .154336)
/rpt_family="AluY45"
repeat_region 154373. .154415
/rpt_family="MER68B"
repeat_region 154461. .154624
/rpt_family="MIR"
repeat_region 154864. .155159
/rpt_family="AluSx"
repeat_region 156920. .157056
/rpt_family="MIR"
repeat_region complement(157057. .157365)
/rpt_family="AluYb8"
repeat_region complement(157368. .157390)
/rpt_family="(TAA)n"
repeat_region complement(157391. .157640)
/rpt_family="AluSg"
repeat_region 157641. .157687
/rpt_family="MIR"
repeat_region complement(158254. .158275)
/rpt_family="(CAAA)n"
repeat_region complement(158288. .158555)
/rpt_family="AluUb"
repeat_region 158792. .159006
/rpt_family="AluSg/x"
repeat_region complement(159894. .160598)
/rpt_family="LIM4"
repeat_region complement(160740. .160773)
/rpt_family="(CATA)n"
repeat_region complement(160993. .161062)
/rpt_family="A_r1ch"
repeat_region complement(161190. .161329)
/rpt_family="MIR"
repeat_region complement(161706. .161780)
/rpt_family="L2"
repeat_region complement(162242. .162312)
/rpt_family="MER46A"
repeat_region 162313. .162615
/rpt_family="MER2"
repeat_region complement(163001. .163293)
/rpt_family="AluY"
repeat_region 165058. .165345
/rpt_family="AluUb"
repeat_region complement(165537. .165724)
/rpt_family="LIMC5"
repeat_region 165939. .166032
```

```
repeat_region      /rpt_family="L2"
                    complement(166033..166330)
repeat_region      /rpt_family="AluSx"
                    166331..166463
repeat_region      /rpt_family="L2"
                    complement(166464..166772)
repeat_region      /rpt_family="AluSg"
                    166773..167134
repeat_region      /rpt_family="L2"
                    167134..167654
repeat_region      /rpt_family="L2"
                    167654..167869
repeat_region      /rpt_family="MIR"
                    complement(167932..168222)
repeat_region      /rpt_family="AluB"
                    complement(168264..168439)
repeat_region      /rpt_family="MER5A"
                    complement(168508..168669)
repeat_region      /rpt_family="MIR"
                    complement(168743..169041)
repeat_region      /rpt_family="AluSc"
                    169191..169298
repeat_region      /rpt_family="L2"
                    169298..169576
BASE COUNT      45170 a 35676 c 38297 g 50887 t
ORIGIN
```

```
Query Match      5.7%; Score 35.2; DB 42; Length 170030;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
```

```
QY 448 tgcgtctcgaggagacgtcctcgagatttaccggaagacaaactggtgtttca 507
Db 22332 TGGTAGCGCCTGTATCCAGACTCAGAGGCTGAGTGAGAAATCGCTTGAAACCA 22391
QY 508 agaatggaagtggtgacctcgagagaccactctccctgaaccacactgtgagt 567
Db 22392 GGAGGTGAGGTTGAGTGAATCAGATCACACCACTGCTGAGCTGGGTGACAGAG 22451
QY 568 atgtcccaagtgcgcagacaccttaaaaaaagaaaaaacaacaaac 623
Db 22452 CAAGACTGCATCTTAACAACAACAAAAAATTAACAACAACAGAAC 22507
```

```
RESULT 15
AGU79997/c      AGU79997      2055 bp      DNA      PLN      06-MAR-1998
LOCUS      Atriplex garrettii male-specific DNA sequence.
DEFINITION      U79997
ACCESSION      92160708
NID
KEYWORDS
SOURCE
ORGANISM      Atriplex garrettii.
                Atriplex garrettii.
                Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
                Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
                Magnoliopsida; Caryophyllales; Chenopodiaceae; Atriplex.
                1 (bases 1 to 2055)
                Ruas,C.F., Fairbanks,D.J., Evans,R.P., Stutz,H.C., Andersen,W.R.
                and Ruas,P.M.
                Male-specific DNA in Atriplex garrettii
                Am. J. Bot. 85, 162-167 (1998)
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
misc_feature
misc_feature
misc_feature
```

```
/note="similar to the Antirrhinum majus (snapdragon) CEN
gene product encoded by GenBank Accession Number S81193
and the Arabidopsis thaliana TFL1 gene product encoded by
GenBank Accession Number U77674; possible pseudogene; the
reading frame is interrupted by repetitive DNA 31 codons
upstream from the stop codon at the point where similarity
to the CEN and TFL1 genes is lost"
BASE COUNT      708 a 311 c 338 g 698 t
ORIGIN
```

```
Query Match      5.6%; Score 35; DB 8; Length 2055;
Best Local Similarity 61.5%; Pred. No. 4;
Matches 56; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 532 agaccacctctccctcagacacacactgtgagtgtgcgaagtgcagacacct 591
Db 1681 ATACATATATTCATCAATATATCCATATATCTAGGTTAACCTAAGTCAATATCTT 1622
QY 592 taataaaaaaagaaaaaacaacaaaaa 622
Db 1621 CAAAAAATAAAAAAAAAAAAAAAAAA 1591
```

```
Search completed: May 14, 1999, 03:35:05
Job time: 5724 sec
```

Oil seed rape cyst
Human retinoid rec

ALIGNMENTS

Search time 129.53 Seconds

(without alignments)
904.854 Million cell updates/sec

.....aaaaaacaaac 623

Issues

Pred. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

[illegible]

OY	449	gctgtctgcggaggatcgtcctccgaagattttaccggagaagacaactggttgcattca	508
Db	378	GCACATCTCGGGGAGACTGCTTGGCCAGCATTTTTFAGGAACAACCACTTCGCTTCAA	437
OY	509	gacatgagagtgtgttcacctcgcgagacaccacactcctccctacgaaccacactgttagta	568
Db	438	GACATGAGAGTGTTGTCTCCTTTGTGGAACACCCTCCTCCTCTTAGAACCGACTGTGCCAGC	497
OY	569	tgtgccaaagtggcagcagacc	589
Db	498	AAGGTCAACCTCTGTGAAGATC	518
RESULT	2		
ID	VJ33362	standard; cDNA to mRNA; 1496 BP.	
AC	VJ33362;		
DT	02-DEC-1998	(first entry)	
NDE	Nucleotide sequence of human beta-OARF065.		
KW	Human; beta-OARF065; stroma cell; antibody; inflammatory;		
KM	cytokine-mediated disease; rheumatism; ulcerative colitis; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	45..1316	
FT		/tag= a	
FT	sig_peptide	/product= "human beta-OARF065 protein"	
FT		45..119	
FT		/tag= b	
FT	mat_peptide	120..1313	
FT		/tag= c	
FT		/transl_except= (pos:711..713, aa= Pro)	
FT		/transl_except= (pos:714..716, aa= Arg)	
PN	M09838304-A1.		
PD	03-SEP-1998.		
PF	26-FEB-1998; J00799.		
PR	27-FEB-1997; JP-043143.		
PI	(ONOY) ONO PHARM CO LTD.		
PI	Fukushima D, Konishi M, Tada H;		
DR	WPI: 98-481205/41.		
DR	P-PADB: W70387.		
PT	Membrane polypeptide expressed by human stroma cells, and antibodies recognising it - for treatment of inflammatory and other cytokine-mediated diseases.		
PT	Dislosure; Pages 40-41; 54pp; Japanese.		
PS	This is the nucleotide sequence of the human beta-OARF065, used in the method of the invention. The process involves the use of peptides expressed by stroma cells, and its antibodies are used for in the prevention and treatment of inflammatory and other cytokine-mediated diseases such as Rheumatism, ulcerative colitis.		
QC	Sequence 1496 BP; 388 A; 360 C; 372 G; 376 T;		

Query Match	52.4%	Score 326.6	DB 1	Length 1496
Best Local Similarity	76.2%	Pred. No. 2.4e-86		
Matches 392	Conservative	0	Mismatches 109	Indels 0
			Gaps 0	
QY	89	caggataataacacgcttctgtgagagccatgacacccaaggtctctactctacacagcagc	148	
Db	18	CAACATTAATATCATTTGATATGAAGAAGATGGCTTTAAAGTCTCTACGTAGAACAAAGAAA	77	
QY	149	gtgccttcgcgtgcacatctctctctactccacctgycatgtgaaagtgtgcgaacc	208	
Db	78	ACGTTTTCACCTCTTTTAGTATTAATTAAGTCATTTTGTCTATGAAGTCACTGTGTAACAA	137	
QY	209	ggaaatltcaggcagcagcaggaattccaagatcatcgatctggaactgttcccttcgcaaacg	268	
Db	138	GGAGACTGTAAACACACAGAATTCAGGGATCGGTGTGGAAATGTGTTCCTCCGCAACACG	197	
QY	269	tctcggaacttgacatggaattgtccaagaatgttgcttcgctctatctggggagagatgcacag	328	
Db	198	TGTGGCGCAGCGATGTGAATTTGTCTAAGGAATATGGCTTCGGCTTAAGGGAGGATGCACAG	257	

QY	329	tgtgtgccccttcagagccggccacccggttttaagaagagacactcggtgtttccaagaagtataagcca	368
Db	258	tgtgtgacgtcccccacacgttcacaaaggttaagaagagacactcggtgtttccaagaatccaagccc	317
QY	389	tgtgtgagactgtgtcgctgtgttaaacgccttccagaagggcgaactgtctacacaccagtgtat	448
Db	318	tctctgacactcgcgacagtggttaaacgccttccagaagggcgaactgtctacacaccagtgtat	377
QY	449	gctgtcttcgaggagactcgtcctgtgccaggtttttacccggaagaccaaactgctgtgttttcaa	508
Db	378	gccactctccggggactcgtcctgtgccaggtttttatggaagacgaaactctgtccgcttttcaa	437
QY	509	gacatctgagttgtgtcccttcggaagaccacactcctccacgaacacacactgttgtatga	568
Db	438	gacatcgagtgctgtgccccttcggaagaccacactcctccacgaacacacactgttgtatga	457
QY	569	tgtgtcccaagtgtgcagcagacc	569
Db	498	aaggtcaacactcgtgaagatc	518

RESULT	3
VII4422	
ID	VII4422 standard; cDNA; 371 BP.
AC	VII4422;
DT	23-JUL-1998 (first entry)
DE	Human secreted protein clone AX92.3 cDNA 5'-end.
KW	Secreted protein; prevention; treatment; gene therapy; ds.
OS	Homo sapiens.
PN	WO9801554-A2.
PD	15-JAN-1998.
PE	07-JUL-1997; U11876.
PR	09-JUL-1996; US-677231.
PA	(GENM) GENETICS INST INC.
PI	Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM,
P1	Merberg D, Racile LA, Spaulding V, Treacy M;
DR	MP1; 98-110230/10.
DR	P-PSDB; W58844.
PT	Secreted proteins and polynucleotides encoding them - useful to
PT	prevent, treat and ameliorate medical conditions
PS	Claim 15; Page 57; 93pp; English.
CC	VII4422-VII4424 encode fragments of a novel secreted protein derived from
CC	clone AX92.3 which was isolated from a human adult testes cDNA library.
CC	The protein can be used to prevent, treat or ameliorate a medical
CC	condition, while the polynucleotides can be used for gene therapy.
SQ	Sequence 371 BP; 83 A; 96 C; 107 G; 82 T;

Query Match	39.9%	Score 248.8	DB 1	Length 371
Best Local Similarity	83.8%	Pred. No. 7.8e-64		
Matches 280	Conservative 0	Mismatches 54	Indels 0	Gaps 0
QY 243	ctggaactgtgtccctctgcgaacaacgtgcygaccctgcgcatgagttgtgccagaatgtg	302		
Db 1	CTGGAACCTGTGTTCCCTCGCAACCAAGTGTGGCCCGCATGAGTTGTCTAAGGAATGTG	60		
QY 303	gcttcgctatgtggagagatgcacagttgttccttcgaagccgcgacccggttcaagaag	362		
Db 61	GCTTCGCTATGGGAGATGCACAGTGTGTGACGTCCGGGTGCACAGGTTCAAGGAGG	120		
QY 363	actgggtttccagaagtgttaagccatgttcggagactgtgcgtgtgtgaacogcttcaga	422		
Db 121	ACTGGGCTTCCAAAATAAGCAAGCCCTGTCTGGACTGCGCAGTGGTGAACCGCTTTCAGA	180		
QY 423	gggcgaactgtccaaacccacccagttgctgtctcggggagactgctctgcacagattttacc	482		
Db 161	AGGCAATTGTTCAACCACCAAGTAGTGCATCTGGGGGAACTGCTTGGCAGATTTTTATA	240		
QY 483	ggaagccaacactgttgttttcaagacatgagttgtgtccctgcggagagaccaactc	542		
Db 241	GGAAACCAAAACTTGTGGGCTTTTCAAAACATGGAATGTGTGCTTGTGAAACCTCTCTC	300		
QY 543	ctccctacgaaccacactgtgagtatgtgcaca	576		


```
DB 301 CTCCTTAGCAACCGCACTGTGCCAGCAAGCTCA 334
|||||
RESULT 4
V11423
ID V11423 standard; cDNA; 181 BP.
AC V11423.
DT 23-JUL-1998 (first entry)
DE Human secreted protein clone AX92_3 cDNA internal fragment.
KW Secreted protein; prevention; treatment; gene therapy; ds.
OS Homo sapiens.
PN M0980154-A2.
PD 15-JAN-1998.
PF 07-JUL-1997; 011876.
PR 09-JUL-1996; US-677231.
PA (GENWAY ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM,
PI Werberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-110230/10.
PR Secreted proteins and polynucleotides encoding them - useful to
PI prevent, treat and ameliorate medical conditions
PS Claim 13: Page 57; 93pp: English.
CC V11422-V11424 encode fragments of a novel secreted protein derived from
CC clone AX92_3 which was isolated from a human fetal brain cDNA library.
CC The protein can be used to prevent, treat or ameliorate a medical
CC condition, while the polynucleotides can be used for gene therapy.
SQ Sequence 181 BP; 43 A; 47 C; 34 G; 57 T;

Query Match 11.3%; Score 70.2; DB 1; Length 181;
Best Local Similarity 85.7%; Pred. No. 1e-11;
Matches 78; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 474 gatttacggagagcccaactggtgtttcagaagatgagtgccctggagag 533
|||
DB 88 GATTTTATGGAAGACGAAGACTGTGCGCTTCAAGACATGATGTGTGGAG 147
|||
OY 534 accacactctccctacagaccacactgtga 564
|||
DB 148 ACCCTCTCTCTCTTACGACCGCACTCTCA 178
|||

RESULT 5
O10388/c
ID O10388 standard; cDNA; 1912 BP.
AC O10388.
DT 11-APR-1991 (first entry)
DE Encodes skin-specific murine Retinoic Acid Receptor-gamma.
KW mouse; retinoic acid receptor; mRAR-gamma; embryogenesis; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 316..1892
FT /tag= a
FT /product= murine RAR-gamma

EP-411323-A.
PD 06-FEB-1991.
PF 29-JUN-1990; 112469.
PR 30-JUN-1989; US-374690.
PR 29-MAR-1990; US-502140.
PA (INRM ) INSERM INST NAT SANTE.
PA (SOUI ) LES LABS SOUIB SA.
PI Chambon P, Kastner P, Krust A, Petkovich M, Zelent A;
PI Leroy P, Mendelsohn C, Staub A;
DR WPI: 91-038271/06.
DR P-PSDB: R105349.
PT Novel human and mouse retinoic acid receptors - encode proteins
PT used to assay for agonists and antagonists
PS Disclosure: Fig 1; 33pp: English.
CC An 11.5 day-old total mouse embryo lambda gt10 cDNA library was
CC screened with labelled human RAR-alpha and -beta probes. Positive
CC clones were isolated, mapped and sequenced. A set of clones was
```

```
CC identified having a lesser degree of homology with either RAR-alpha
CC or -beta than murine RAR-alpha or mRAR-beta clones, respectively.
CC The cDNA-deduced amino acid sequence of this third set of clones
CC was obviously related to mRAR-alpha and mRAR-beta. This new member
CC of the mouse RAR subfamily was designated RAR-gamma. The
CC mRAR-gamma cDNA clone was used to clone 7 different murine
CC RAR-gamma cDNA isoforms, generated by alternative splicing of at
CC least 7 exons. The isoforms were found to differ in their 5'-UTR
CC sequences and in their N-terminal A region, which is known to be
CC important for differential trans-activation by other nuclear
CC receptors.
CC See also R10547-8, Q10389 and Q10405-8.
SQ Sequence 1912 BP; 427 A; 597 C; 534 G; 354 T;

Query Match 5.6%; Score 35; DB 1; Length 1912;
Best Local Similarity 47.5%; Pred. No. 0.53;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

OY 207 ccgagatgcagcagcagaatccaagatcgaatctggaactgtctcttcaaac 266
|||
DB 233 CCTGGCCTGGAGAGCTCCGTCGCCACCTGCTGCTGCCAGTGGGTTTCCAGGATC 174
|||
OY 267 agtgcgacctgcatggaatgttccaagaaatgtgcttgctatggagagatgcac 326
|||
DB 173 CCTTGGCCTCTGCAGAGATCCCTCTACATACGTGGCCAGCTGAGGTGAGACCTG 114
|||
OY 327 agtgcgtgcctgcagccgacacgcgttcaaggaagactgggttccagaagtgaac 386
|||
DB 113 GTGGAACCTGTGAGAGGAGGAGGCGGGGAGAGGAGGCCCGCAGAGCCCGAGGTCCCGGC 54
|||
OY 387 catgtcggaactgtgcgtgtgaacgcgttccagaagg 425
|||
DB 53 GTCGGGCGGTCTCGGGGATGAGACCGCGGCTGGGGGG 15
|||

RESULT 6
T05202/c
ID T05202 standard; cDNA; 1913 BP.
AC T05202.
DT 13-JUN-1996 (first entry)
DE Murine retinoic acid receptor gamma cDNA.
KW Retinoic acid receptor; mouse; RAR; RAR-gamma; transcription factor;
KW skin; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 316..1692
FT /tag= a
FT /product= retinoic acid receptor gamma protein
FT misc_feature 316..498
FT /tag= b
FT /note= "hRAR-gamma A region"
FT misc_feature 499..582
FT /tag= c
FT /note= "hRAR-gamma B region"
FT misc_feature 583..780
FT /tag= d
FT /note= "hRAR-gamma C region"
FT misc_feature 781..918
FT /tag= e
FT /note= "hRAR-gamma D region"
FT misc_feature 919..1578
FT /tag= f
FT /note= "hRAR-gamma E region"
FT misc_feature 1579..1689
FT /tag= g
FT /note= "hRAR-gamma F region"

EP-683227-A1.
PD 22-NOV-1995.
PF 29-JUN-1990; 110924.
PR 30-JUN-1989; US-374690.
PR 29-MAR-1990; US-502140.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
```

PA (BRIM) INST RECH SOUTIB.
 PA (BRIM) LES LAB SOUTIB SA.
 PI Chamoun P, Kastner P, Krust A, Leroy P, Mendelsohn C;
 PI Petkovich M, Stauda, Zelent A;
 DR MPI; 95-394349/51.
 DR P-PSDB; R84726.
 PT New cDNA encoding the human retinoic acid receptor gamma - useful
 PT for identifying agonists and antagonists of RAR-gamma and detecting
 PT expression in normal and diseased tissue
 PS Example 1: Fig 1: 34pp: English
 CC This sequence represents the cDNA encoding the mouse retinoic acid
 CC receptor gamma (RAR-gamma). This sequence was used in the isolation of
 CC the human RAR-gamma sequence. RAR-gamma is a retinoic acid inducible
 CC transcription factor, and is expressed specifically in the skin where it
 CC is the predominant RAR. The C and E regions of the protein are the most
 CC highly conserved regions through the RAR's of a species. The C region is
 CC the DNA binding domain, and the E region is the ligand binding domain.
 CC The B region is also well conserved within species. The N and C terminal
 CC regions of segment D are conserved, whereas the central segment of this
 CC region is not. The D region is thought to act as a hinge region.
 CC Regions A and F are not conserved within a given species. The A, B, D
 CC and F regions are conserved across species, suggesting that they have
 CC specific functions, but are not required for the RAR's specific
 CC physiological roles. The protein is used to detect competitive and
 CC non-competitive agonists and antagonists. This sequence can be used to
 CC detect receptor expression in normal and diseased tissues and cultures.
 SQ Sequence 1913 BP; 427 A; 600 C; 531 G; 355 T;

Query Match 5.6%; Score 35; DB 1; Length 1913;
 Best Local Similarity 47.5%; Pred. No. 0.53;
 Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 207 ccgagagatcgcagcgagcgaattcaagatgatctgaactgtctctgcaaac 266
 DB 233 cctggccttgagagcctcctgcccacccctcctgctgctcctgctccaaagatc 174
 QY 267 agtgcgagcctgcagatgtgtccaaagaatgtgctcgcgtcgaaggagatgcac 326
 DB 173 cctctgagccttcgacagagatccctctctacatctggtgagcagctgagagagcctg 114
 QY 327 agtgtgtccctcagcgccgacacggttcacagaagaatcgtgggttcacagaagtgtgaagc 386
 DB 113 gtgcgaactcctgagagcgagcgagggaaggaagcgcacagaccccgaggtcccgccg 54
 QY 387 catgtgagcagctgtgctgtgaaccgcttcagaagg 425
 DB 53 gtgcggcgtctcggagatgacacccgctggtggggg 15

RESULT 7
 V39359 standard; cDNA; 1013 BP.
 AC V39359;
 DE 21-SEP-1998 (first entry)
 DT Humanised anti-HM1.24 antibody encoding cDNA.
 KM Mouse; human; humanised; anti-HM1.24 antibody; myeloma; FR; CDR;
 KW framework region; complementarity determining region; antigenicity; ss.
 OS Synthetic.
 OS Mus sp.
 OS Homo sapiens.
 FH Homo sapiens.
 FT Key Location/Qualifiers
 FT CDS 23..565
 FT /tag= a
 PN MO9814560-A1.
 PD 09-APR-1998.
 PF 03-OCT-1997; J03553.
 PR 04-OCT-1996; JP-264756.
 PA (CHUS) CHUGAI SEIYAKU KK.
 PI Koishihara Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M,
 PI Yoshimura Y;
 DR MPI; 98-286421/25.
 DR P-PSDB; W62207.

PT Humanised anti-HM1.24 antibody - for treatment of myeloma
 PS Claim 81: Page 151-153; 210pp: Japanese.
 CC A humanised anti-HM1.24 antibody has been developed which comprises
 CC human L and H chain C regions, and L and/or H chain V regions
 CC containing material originating in mouse anti-HM1.24 antibody. The V
 CC regions contain framework (FR) regions of human origin and
 CC complementarity determining regions (CDR) of mouse origin, leading to
 CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and
 CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the
 CC L chain V region are derived from human subtype HSG1 (e.g. from human
 CC antibody RB1) and the FR regions of the H chain V region are derived
 CC from human subtype HSG1 (e.g. FR1-3 from human antibody Hg3 and FR4
 CC from human antibody JH6). The present sequence encodes an antibody
 CC polypeptide from the present invention. The antibodies are used for
 CC the treatment of myeloma, especially by injection, intravenously,
 CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000
 CC (especially 5-100) mg/kg body weight. The humanised antibody has low
 CC antigenicity and is therefore effective therapeutically in humans.
 SQ Sequence 1013 BP; 232 A; 254 C; 314 G; 213 T;

Query Match 5.5%; Score 34.2; DB 1; Length 1013;
 Best Local Similarity 83.0%; Pred. No. 0.71;
 Matches 39; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 576 agtgcgagcagaccttaaaaaaaaaaagaaacaaacaaaaa 622
 DB 953 agcgagagcacccttataaaaaaaaaaaaaaaaaaaaaa 999

RESULT 8
 V07579 standard; cDNA; 1013 BP.
 AC V07579;
 DE 20-NOV-1998 (first entry)
 DE Protein recognised by anti-human HM1.24 antibody nucleotide sequence.
 KW ss; cytotoxic antibody; anti-human HM1.24; lymphocytic tumours;
 KW pre-B lymphoma; Burkitt's lymphoma; T-cell tumour; B-cell tumour.
 OS Homo sapiens
 FH Homo sapiens
 FT Key Location/Qualifiers
 FT CDS 23..565
 FT /tag= a
 FT /product= "Cytotoxic protein"

PD MO9835698-A1.
 PD 12-FEB-1998; J00568.
 PR 12-FEB-1997; JP-041410.
 PA (CHUS) CHUGAI SEIYAKU KK.
 PI Koishihara Y, Yoshimura Y;
 DR MPI; 98-456869/39.
 DR P-PSDB; W65771.
 PT Treatment of lymphocytic tumours using cytotoxic antibody - binding
 PT to specific antigen such as HM1.24 and effective against T-cell
 PT tumours and B-cell tumours other than myeloma
 PS Claim 1: Page 44-45; 82pp: Japanese.
 CC The protein having the nucleotide sequence below is bound specifically
 CC by a cytotoxic antibody which can be used in the treatment of lymphocytic
 CC tumours, including T-cell tumours and B-cell tumours other than myeloma.
 CC The antibody is preferably monoclonal and has ADCC or CDC type
 CC cytotoxicity. It may be chimeric or humanised, and preferably contains
 CC a human antibody constant region C gamma (such as C gamma 1 or C
 CC gamma 3). A preferred antibody is an anti-human HM1.24 antibody or an
 CC antibody which binds to an epitope recognising anti-human HM1.24
 CC antibody. The cytotoxic antibody is useful in the treatment of
 CC lymphocytic tumours such as acute or chronic B lymphocytic leukemia,
 CC pre-B lymphoma, Burkitt's lymphoma, or acute or chronic T lymphocytic
 CC leukaemia.
 SQ Sequence 1013 BP; 232 A; 257 C; 311 G; 213 T;

Query Match 5.5%; Score 34.2; DB 1; Length 1013;
 Best Local Similarity 83.0%; Pred. No. 0.71;
 Matches 39; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 576 agtgcagcagcagccttaaaaaaagaaaaaacaacaaaaa 622
 DB 953 AGGGAGACGACACCTTAATAAAAAAAAAAAAAAAAAAAAA 999

RESULT 9

V59067
 ID V59067 standard: cDNA; 2634 BP.

AC V59067;

DT 02-FEB-1999 (first entry)

DE Bax inhibitor Bi-1 cDNA

KW Bax inhibitor; Bi-1; human; apoptosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 73..786

FT CDS /tag= a

PN MO9840397-A1.

PD 17-SEP-1998.

PF 13-MAR-1998; 005015.

PR 14-MAR-1997; 05-818514.

PA (BURN-) BURNHAM INST.

PI Reed JC, Xu Q;

DR WPI: 98-531519/45.

P-PSDB: W73136.

PT Bax inhibitor proteins, Bi-1 and Bi-2 - useful e.g. to modulate

cellular apoptotic activity or identify agents altering Bi-1 or Bi-2

binding which can modulate apoptotic activity

Claim 2: Page 61-63; 80pp; English.

CC This cDNA clone codes for an inhibitor protein, termed Bi-1 (see

W73136), of the pro-apoptotic protein Bax. Nucleic acids encoding

Bi-1 and Bi-2 (see V59068) were identified by suppression of

Bax-induced death of yeast cells transformed to express human Bax.

CC A human HepG2 cDNA library was used for library screening. The

CC invention provides vectors, optionally expression or viral vectors,

CC containing Bi nucleic acids, and host cells containing these

CC vectors. The nucleic acids encoding Bi-1/Bi-2 can be used to

CC increase expression of these proteins in cells, or antisense

CC molecules prepared from them used to decrease expression. In

CC these ways, cellular apoptotic activity may be modulated (claimed).

CC The nucleic acids and complementary sequences are also useful as

CC probes to detect Bi-encoding nucleic acid molecules in samples.

CC Sequence 2634 BP; 632 A; 626 C; 597 G; 779 T;

SO

Query Match 5.4%; Score 33.8; DB 1; Length 2634;

Best Local Similarity 77.4%; Pred. No. 1.4;

Matches 41; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 570 gtgcacagtgagcagccttaaaaaaagaaaaaacaacaaaaa 622

DB 2579 GTGCCAAGATGACAGATATTAATAAAAAAAAAAAAAAAAAAAAA 2631

RESULT 10

V59114

ID V59114 standard: cDNA; 1016 BP.

AC V59114;

DT 14-DEC-1998 (first entry)

DE Nucleotide sequence of protein bound by Anti-HM1.24 antibody.

KW ss; Anti-HM1.24; antibody; lymphocyte activation inhibitor;

KW autoimmune disease; organ transplant; allergy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 23..565

FT CDS /tag= a

FT CDS /product= "Protein bound by Anti-HM1.24"

PN MO9837913-A1.

PD 03-SEP-1998.

PF 27-FEB-1998; J00831.

PR 28-FEB-1997; JP-045663.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Koshinara Y;

QY 581 cagcagaccccttaaaaaaagaaaaaacaacaaaaa 622

DB 3582 GTGGCAGCGATGATCAGAGACCTCTCTAGCCCTATACCTGTCAATAAATCTTGTTA 3641

RESULT 11

T72107

ID T72107 standard: cDNA to mRNA; 3692 BP.

AC T72107;

DT 13-FEB-1998 (first entry)

DE Rat semaphorin 2 gene.

KW Semaphorin 2; central nerve extension; rat; human; inhibitor;

KW central nerve regeneration promoter; ss.

OS Rattus norvegicus.

FH Key Location/Qualifiers

FT 5'UTR 1..18

FT CDS /tag= a

FT CDS 19..2662

FT CDS /tag= b

FT CDS 2683..3653

FT CDS /tag= c

FT CDS 3654..3692

FT CDS /tag= d

PN MO9720928-A1.

PD 12-JUN-1997.

PF 02-DEC-1996; J03517.

PR 31-OCT-1996; JP-307205.

PR 06-DEC-1995; JP-345187.

PA (SUMI) SUMITOMO PHARM CO LTD.

PI Kikuchi K, Kimura T;

DR WPI: 97-319775/29.

P-PSDB: W19856.

PT Semaphorin 2 and the gene encoding it - also inhibitors of its

action which can be used as promoters of central nerve regeneration

Claim 1: Page 61-63; 10pp; Japanese.

CC This sequence represents the rat semaphorin 2 gene. The protein encoded

CC by this sequence, and the human semaphorin 2 protein (see W19857) can be

CC used for screening possible candidates for activity as semaphorin 2

CC inhibitors. Proteins containing all or part of the semaphorin 2 sequence

CC are useful as central nerve extension inhibitors. Semaphorin 2 inhibitors

CC identified by screening can be used as promoters of central nerve

CC regeneration.

CC Sequence 3692 BP; 689 A; 1126 C; 1122 G; 755 T;

SO

Query Match 5.3%; Score 33.2; DB 1; Length 3692;

Best Local Similarity 57.8%; Pred. No. 2.3;

Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 521 gtgcctgcagagaccacccctccctacgaacacacactgtgagtgccaagtgg 580

DB 3582 GTGGCAGCGATGATCAGAGACCTCTCTAGCCCTATACCTGTCAATAAATCTTGTTA 3641

DB 3642 CATCCACGCCCAAAAAAAAAAAAAAAAAAAAAA 3683

RESULT 12

ID T89343 standard; cDNA; 2083 BP.

AC T89343.

DT 11-MAR-1998 (first entry)

DE Human p62 cDNA #1.

KW p62; cytoplasmic; T cell; B cell; development; activation; modulation; cellular response; cell proliferation; autoimmune disease; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 67..1389

FT /*tag= a

FT /product= p62

PN MO9722255-A1.

PD 26-JUN-1997.

PR 11-DEC-1996; U19944.

PS 19-DEC-1995; US-574959.

PA (DAND) DANA FARMER CANCER INST INC.

PI Jung I, Shin J, Strominger JL, Vadmud RK;

DR WPI: 97-341351/31.

DR P-PSDB: W31182.

PT cDNA encoding p62 and p160 and corresponding proteins - used in the treatment of autoimmune disease and for T and B cell proliferation, e.g. for treatment of tumours

PS Claim 4: Fig 1: 175pp; English.

CC This cDNA sequence encodes a novel p62 cytoplasmic polypeptide which is expressed in a variety of tissues including heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. This polypeptide is capable of modulating T or B cell development and/or T or B cell activation e.g. by modulation of cell activity. It is also capable of modulating degradation of cellular proteins e.g. cell cycle regulatory proteins stimulating expression of cell cycle dependent kinase inhibitors and arresting cell cycle progression at specific boundaries to thereby modulate cell proliferation. As p62 acts to boost B cell response and may be used to treat disorders where this is beneficial, e.g. infections by pathogenic microorganisms, e.g. bacteria, viruses and protozoans. p62 can be used to expand T cell populations for treating infectious diseases or cancer, e.g. the resulting cells may be transduced to render them resistant to HIV infection. Inhibitors of p62 can be used to reduce B or T cell responses and may be used to treat a variety of autoimmune diseases, e.g. diabetes mellitus, arthritis, multiple sclerosis allergic reactions, Crohn's diseases etc.

CC Sequence 2083 BP; 447 A; 588 C; 608 G; 440 T;

SO

Query Match 5.3%; Score 33; DB 1; Length 2083;

Best Local Similarity 79.6%; Pred. No. 2.1;

Matches 39; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 575 aagtgagcagcagccttcaaaaaaagaacaaacaaac 623

DB 2027 AATGTGACATGACACTTAAACCAATMAAAAAAAAAAAAAAAAAAAC 2075

RESULT 13

ID T73502 standard; cDNA to mRNA; 1868 BP.

AC T73502.

DT 16-SEP-1997 (revised)

DT 26-AUG-1997 (first entry)

DE Poplar 1-aminocyclopropane-1-carboxylic acid synthase gene PNACCS2.

KW Ozone; induction; exposure; resistance; transgenic plant; ACC;

KW 1-aminocyclopropane-1-carboxylic acid biosynthesis; ds.

OS Populus nigra.

PH Key Location/Qualifiers

FT CDS 119..1579

FT /*tag= a

FT /product= PNACCS2

PN J09075088-A.

PD 25-MAR-1997.

PF 07-SEP-1995; 254510.

PR 07-SEP-1995; JP-254510.

PA (TOYT) TOYOTA JIDOSHA KK.

DR WPI: 97-239270/22.

DR P-PSDB: W21755.

PT Ozone-induced 1-amino:cyclopropane-1-carboxylic acid synthetic enzyme genes - from poplar tree, are useful for generating ozone-resistant trees and pollution clean-up trees

PS Claim 2: Pages 9-11; 12pp; Japanese.

CC This sequence, designated PNACCS2, is a 1-aminocyclopropane-1-carboxylic acid (ACC) synthase gene isolated from poplar trees which had been exposed to ozone. Exposure conditions were: 0.1 ppm ozone for 6 hours after one hour acclimation at 25 degrees C, 70 % relative humidity, 30 k lux and air current 30 cm/second. This gene will be useful for breeding air pollutant ozone-resistant trees, especially poplar.

CC (Note: Record has been revised in order to correct the Organism Source field).

CC Sequence 1868 BP; 590 A; 366 C; 405 G; 507 T;

SO

Query Match 5.3%; Score 32.6; DB 1; Length 1868;

Best Local Similarity 67.6%; Pred. No. 2.3;

Matches 46; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 555 cacactgagtgatgtgtccagatgacgagccttcaaaaaaagaacaaac 614

DB 1774 CATTTGAAATTCAGATGATATTTGTTCTCTCTTTAAAAAAAAAAAAAAAA 1833

OY 615 aacaaaaa 622

DB 1834 AAAAAAAAA 1841

RESULT 14

ID 054832 standard; DNA; 1227 BP.

AC 054832.

DT 19-JUL-1994 (first entry)

DE Sequence of clone W225 encoding mouse gonadotropin-releasing hormone receptor (GnRH-R).

KW Gonadotropin-releasing hormone receptor; GnRH; diagnosis; contraceptive; reproductive disorders; ss.

OS Mus musculus.

PH Key Location/Qualifiers

FT CDS 43..1023

FT /*tag= a

PN M09400590-A.

PD 06-JAN-1994.

PF 22-JUN-1993; U05965.

PR 23-JUN-1993; US-904072.

PR 21-JUN-1993; US-080386.

PA (UYN) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

PI Sealton SC;

DR WPI: 94-026225/03.

DR P-PSDB: R47164.

PT Gonadotropin-releasing hormone receptor genes and proteins - for expression of GnRH and screening and identifying GnRH (ant)agonists, for diagnosis and therapy of reproductive disorders and for contraception

PS Claim 8: Page 44-46; 73pp; English.

CC The longest ORF of 054832 encodes a 327 AA protein of about 37,000 MW (R47164). The product is murine GnRH-R. Three consensus N-linked glycosylation sites are present, two in the N-terminus and one in the first extracellular loop. Hydrophobicity analysis of the deduced protein reveals seven stretches of highly hydrophobic AAs with 20-30% sequence similarity to other G-protein receptors, with the highest degree of homology to the interleukin-8 receptor.

CC Sequence 1227 BP; 337 A; 308 C; 246 G; 336 T;

SO

Query Match 5.2%; Score 32.6; DB 1; Length 1227;

Best Local Similarity 80.9%; Pred. No. 2.2;

Matches	38;	Conservative	0;	Mismatches	9;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY	576	agtggcagcagaccc	ttaaaaaaaaaa	aaga	aaaaaa	acaacacaaa	622
Db	1174	AGCTTCAGAAGACCTTCAAAAACAAAAA	AAAAAAAAAAAAAA	AAAAA	AAAAA	AAAAA	1220

RESULT 15

ID	T37306	standard; cDNA; 1227 BP.
AC	T37306;	
DT	04-DEC-1996	(first entry)
DE	Murine gonadotropin-releasing hormone receptor cDNA clone WZ25.	
KW	Gonadotropin-releasing hormone receptor; GnRH-R; G-protein receptor	
KW	signal transduction; reproduction; contraception; diagnosis;	
KW	antisense; ribozyme; gene therapy; transgenic animal; ss.	
MS	Mus sp.	
FS		
FT	Key	Location/Qualifiers
FT	cds	43..1020
FT		/tag- a
PN	W09625423-A1.	
PD	22-AUG-1996.	
PF	26-JAN-1996; U01034.	
PR	17-FEB-1995; US-390000.	
PA	(MOUN) MOUNT SINAI SCHOOL MEDICINE.	
PI	Sealion SC;	
DR	WPI; 96-393334/39.	
DR	P-PSDB; W039395.	
PT	identifying modulators of gonadotropin-releasing hormone receptor -	
PT	including new anti-sense oligo:nucleotide(s) and antibodies, useful	
PT	e.g. for contraception or diagnosis and treatment of reproductive	
PT	disorders	
PS	Example 6; Fig 3; 76pp; English.	
CC	A cDNA clone (T37306), designated WZ25, codes for murine	
CC	gonadotrophin-releasing hormone receptor (GnRH-R) (W03995),	
CC	a G-protein receptor involved in the reproductive system. It	
CC	was isolated from the mouse gonadotropin cell line alpha-T3-1 by	
CC	PCR amplification using primers (see also T37302-03) based on	
CC	conserved regions of G-protein receptor, use of an isolated clone	
CC	(WZ7) to screen an alpha-T3-1 cDNA library, and by hybrid-arrest	
CC	and expression assays. The cDNA was used to isolate human GnRH-R	
CC	cDNA (T37307). It may also be used to produce recombinant GnRH-R	
CC	in transformed host cells, or to breed transgenic animal models	
CC	for evaluation of GnRH analogues in vivo.	
CC	Sequence 1227 BP; 337 A; 308 C; 246 G; 336 T;	

Query Match	Score	DB	Length
5.28;	32.6;	DB 1;	1227;

Best Local Similarity	80.3%;	Pred No. 2.2;							
Matches	38;	Conservative	0;	Mismatches	9;	Indels	0;	Gaps	0;

QY	576	agtgcagcagaccc	ttaaaaaaaaa	agaaaaaa	aacacacaaaa	622
Db	1174	AGCTTCAGAGACCTT	CAAAAA	CAAAAA	AAAAAAAAAAAA	1220

Search completed: May 14, 1999, 11:42:35
Job time: 5264 sec

This Page blank (uspio)

Db	236	TGCAAAAGTGGCGACNCTGGCATGAGTGTGTCCAAAGATGTGGCTTCGGCTAATGGGAG	295
Qy	320	gatcacagtgctgtccctcgagccgcgacccggttccaagaagactgggtttcagaag	379
Db	236	GATGACAGATGTGTGCTCCCTGAGCGCCGACCGCGTTCAGGAAGACTGGGCTTCAGAAG	355
Qy	380	tgt 382	
Db	356	tgt 358	
RESULT	3		
LOCUS	AA575933/c		
DEFINITION	AA575933 520 bp mRNA EST 09-SEP-1997 nm6c10.s1 NCI_CGAP_Br3 Homo sapiens cDNA clone IMAGE:1072242 similar to SW:COX1_HUMAN P00395 CYTOCHROME C OXIDASE POLYPEPTIDE I ; mRNA sequence.		
ACCESSION	AA575933		
NID	G2350448		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 520) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)		
JOURNAL			
COMMENT	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emert-Buck, M.D., Ph.D. CDNA Library Preparation: Stratagene, Inc. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: www.bio.lnl.gov/dbtrp/image/image.html		
FEATURES	Insert Length: 715 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amersham High quality sequence stop: 322. Location/Qualifiers 1..520 /organism="Homo sapiens" /note="Organ: breast; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Ductal breast tumor. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average Insert size: 0.9 kb." /db_xref="taxon:9606" /clone="IMAGE:1072242" /clone_11b="NCI_CGAP_Br3" /sex="female" /tissue_type="breast tumor" /lab_host="SOLR (kanamycin resistant)" /1..>520 BASE COUNT 139 a 86 c 133 g 162 t		
ORIGIN	MRNA		
BASE COUNT	139 a 86 c 133 g 162 t		
ORIGIN			
Query Match	7.3%; Score 45.6; DB 22; Length 520;		
Best Local Similarity	52.0%; Pred. No. 0.0095;		
Matches 102; Conservative	0; Mismatches 94; Indels 0; Gaps 0;		
Qy	425	gccacactgctcacacacacagatgctgtcgtcgaggagctgctccagatattacg	484
Db	196	GCTGACTGCCCACTCCACGGAAGCAATATATAATGATCTCTGACGTCTTGAGCC	137
Qy	485	aagacacactgtgtgttttcaagacatgagagtgctgtccctcggaaccactct	544

DB	136	TGAGTCATCTTTTCTTTTCACCGTAGGTCGCTACTGCGCATTTGTATTATGCAACTCAT	77
QY	545	cccttcgacacacacacatgtagtgaatgctgcaagtgagcgacacaccttaaaaaaaga	604
DB	76	CACGTGACATCGTACTACTAACGACGACAGTACTACTGTTGATGCTTCCACTTAAAAAAA	17
QY	605	aaaaaaacaacaaca 620	
DB	16	AAAAAAAAAAAAAAAAA 1	
RESULT	4		
LOCUS	T92700/c		
DEFINITION	T92700	321 bp	mRNA EST 22-MAR-1995
ACCESSION	Y26901.1		
NID	ye6901.1		
KEYWORDS	gb: X13546.ra1		
SOURCE	Human HMG-17 gene for non-histone chromosomal protein (HUMAN);		
ORGANISM	Human clone=118896 library=Stratagene lung (#937210)		
REFERENCE	vector=pbuuescript SK- host=SOB cells (kanamycin resistant)		
AUTHORS	primer=Mi3RPI Rsite1-EcoRI Rsite2-XhoI Normal lung tissue from a 72 year old male. Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence: 5'-GAATTCGGCAGAG-3'; 3' adaptor sequence: 5'-CTCGAGTTTCTTTTCTTTTCTTTT-3'.		
TITLE	Homo sapiens		
JOURNAL	Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
COMMENT	1 (bases 1 to 321) Hallier,M., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaaks,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R. WashU-Merck EST Project Unpublished (1995)		
FEATURES	Source		
BASE COUNT	94 a	43 c	80 g 100 t 4 others
ORIGIN			
Query Match	6.3%	Score 39.2:	DB 10; Length 321;
Best Local Similarity	65.9%	Pred. No. 0.5;	
Matches	56;	Conservative 0;	Mismatches 29; Indels 0; Gaps 0;
QY	535	ccctactctctcctcctcgcacacacatgtagtgaatgctgcaagtgagcgacaccttaa	594
DB	261	CCCCCTTCCCCCCCCAACACAAATGAAGTGTGTGCTTAACAACATAGCTTTTA	222
QY	595	aaaaaaagaaaaaaacaaca 619	
DB	221	AAAAAAAAAAAAAAAAACTTAAACAAA 157	
RESULT	5		
AA177203			

LOCUS	AAL77203	417 bp	mRNA	EST
DEFINITION	mtl3e04.r Soares mouse 3NDMS Mus musculus cDNA clone 620958 5' ,			
ACCESSION	AAL77203 mRNA sequence.			
NID	91759651			
KEYWORDS	EST .			
SOURCE	house mouse. Mus musculus			
ORGANISM	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Muinae;			
REFERENCE	Mus. 1 (bases 1 to 417) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gaisel,S., Kliccaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R. The WashU-HHMI Mouse EST Project Unpublished (1996)			
TITLE				
JOURNAL				
COMMENT	<p>Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mousee@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:381782 Seq primer: -28M13 rev2 from Amershams High quality sequence stop: 384. Location/Qualifiers 1..417 /organism="Mus musculus" /strain="C57BL/6J" /note="vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAGTCGTGAAGTGGAGCGCCGCCTGTATTTTTTTTTTTTTTTT 3'] : double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo." /db_xref="taxon:10090" /clone="620958" /clone_lib="Soares mouse 3NDMS" /sex="male" /tissue_type="Spleen" /dev_stage="4 weeks" /lab_host="DH10B" <1..>417</p>			
BASE COUNT	129 a	107 c	98 g	83 t
ORIGIN	mRNA			
	<1..>417			
Query Match	6.3%	Score 39.2:	DB 16:	Length 417:
Best Local Similarity	69.7%:	Pred. No.0.51:		
Matches	53:	Conservative	0:	Mismatches 23: Indels 0: Gaps 0.0:
OY	547 ctacgaaccacctgtagtgagtcgcaggcgagcacactttaaaaaaaagaagaa	606		
Dd	335 CTACCAAGTCGCCACTGTGGTGGAGAATAAAAGTGCTTCTTAATAAAAAAAA	394		
OY	607 aaaaacaacaacaa 622			
Dd	395 AAAAAAAAAAAAAA 410			

[illegible]

REFERENCE

Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 306)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T., Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Mareston,R.

The WashU-HHMI Mouse EST Project
unpublished (1996)

TITLE

MashU-HHMI Mouse EST Project

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project
MashU-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGJ:635204

Putative full length read
vector to vector length is 702
Seq primer: -28m13 rev2 Et from Amersham
High quality sequence stop: 95.

FEATURES

source Location/Qualifiers

- 1..306 /organism="Mus musculus"
/strain="FVB/N"
/note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker. Site.1: EcoRI; Site.2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gy's. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'TGTACGATCTGAAGTGAGCGCCGCCCTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (TAATCGATCCTGG), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT3 vector. Library constructed by Bob Barstead."
/db_xref="taxon:10090"
/clone_id="Baxton:10090"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"

BASE COUNT

177 a 45 C 39 g 45 t

ORIGIN

```
Query Match          6.1%; Score 38.2; DB 26; Length 306;  
Best Local Similarity 63.7%; Pred.No. 0.93;  
Matches    58; Conservative   0; Mismatches     33; Indels   0; Gaps  0;
```

Oy 532 agaccacctcctccctgcgaaccacactgtagatgtgccaaagtggcacagaccctt 591
||||| | | | | ||||| | | | | ||||| | | | | ||||| | | | |
Db 131 AGACGCCTCAACTGCATGATGAGAATTAACTACCTGCTGCTCATATAAAGTTCACAAGACATTTA 190

Oy 592 taataaaaaagaataaaaaaacacaacaaa 622
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 TAAAAAAAIAAAAAAAAAAAAAAAAAAAA 221

RESULT

LOCUS	A1249880	255 bp	mRNA	EST	05-NOV-1998
DEFINITION	gx03e09.x1 NCL_GAP_Lym12 Homo sapiens CDNA clone IMAGE:2000296 3'				
ACCESION	Similar to contains Alu repetitive element.; mRNA sequence.				
NID	A1249880				
KEYWORDS	g3846409				
SOURCE	EST.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
	Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 255)				

AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL COMMENT		Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov Unknown library type Seq primer: -40bp from GlbcO. Location/Qualifiers 1..255
FEATURES source		/organism="Homo sapiens" /note=Organ: Lymph node; Vector: pCMV-Sport6; Site_1: Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.25 kb. Life technologies catalog #: 11547-015" /db_xref="taxon:9606" /clone_lhb="NCI-CGAP_Lym12" /tissue_type="Lymphoma, follicular mixed small and large cell"
BASE COUNT ORIGIN	43 a 60 c 54 g 98 t	/lab_host="DH10B"
Query Match Best Local Similarity 57.8%; Pred.No.1.3;	Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;	
Oy 506 caagacatgtagtgctgtgcctcgaggagaccacctcctccctaagaacaactgtagt	565	
Dd 116 CAAAAGGGCCGGGTTCACGTGCCTCAGATGCCGAGATCCCCCCCCTGCCCTCACACTGGTGACAG	57	
Oy 566 tgatttcgaagtggcacgacacctttaaaaaaaagaataaaaacaaacaaa	621	
Dd 56 AATGAGACCCTGCTCAAAAAAAAAAAAAAAAAAAAAA	1	
RESULT 9 LOCUS C62130 360 bp mRNA EST 29-AUG-1997 DEFINITION C.elegans cdna clone yk270b5 : 5' end, single read, mRNA sequence. ACCESSION C62130 NID g2420835 KEYWORDS ESR: EST(expressed sequence tag). \ Caenorhabditis elegans (Strain:CBL489 him-8(e1489)) varied hermaphrodite, male whole animal CDNA to mRNA, clone_lhb:yuj1 Kohara unpublished cdna library clone:yk270b5. ORGANISM Caenorhabditis elegans Eukaryotes; Metazoa; Nematozoa; Secernenteae; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea; Rhabditidae; Pelodermeineae; Caenorhabditis. REFERENCE 1 (sites) Kohara,Y., Motohashi,T., Tabara,H., Shin-i.T., Watanabe,H., Sano,M., Miyata,A., Ohba,T., Mitani.Y., Uesugi.H., Sugita,I., Obana.M., Sugimoto.A., Iida,K. and Nishigaki,A. Expression map of the C.elegans genome JOURNAL TITLE Direct Submission submitted (22-JUL-1997) to the DDBJ/EMBL/Genbank databases. YUJI Kohara, National Institute of Genetics, Gene Network Lab: Yata 1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp, Tel:+81-559-81-6854, Fax:+81-559-81-6855) FEATURES source Location/Qualifiers 1..360 /organism="Caenorhabditis elegans" /strain="CB1489 him-8(e1489)" /db_xref="taxon:6239" /clone_yk270b5" /clone_lhb="yuji kohara unpublished cdna library"		

```
/dev_stage="varied"
/sex="hermaphrodite, male"
/tissue_type="whole animal"

BASE COUNT      145 a      65 c      55 g      88 t      7 others
ORIGIN

Query Match      6.0%: Score 37.4; DB 22; Length 360;
Best Local Similarity 63.6%; Pred. No. 1.6;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 535 cccacccctccctacgaaccacactgtgagtgcagagtcgacagaccttaaa 594
      |||||  |||  |||  |||||  |||||  |||  |||  |||  |||  |||
Db 165 CTCACCCCGCACCCACACACATCGAGTGTATTAATAAANAATTGATTTTAA 224

QY 595 aaaaaaagaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 225 AAAAAAAAAAAAAAAAAAAAAA 252

RESULT 10
AA837213/c      299 bp      mRNA      EST      31-MAY-1998
LOCUS      o024f11.s1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1368909,
DEFINITION      mRNA sequence.
ACCESSION      AA837213
NID      q2912412
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 299)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldio, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMIL at:
www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 721 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 289.
Location/Qualifiers
1..299
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IGD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). CDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTCCAACTCAAGTGGAGCGGCCCTATTTTTTTTTTTTTTTTTT-
3']. Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldio."
/db_xref="taxon:9606"
/clone="IMAGE:1368909"
/clone_11b="NCI_CGAP_GCB1"
```

```
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT      66 a      78 c      65 g      90 t
ORIGIN

Query Match      5.9%: Score 37; DB 27; Length 299;
Best Local Similarity 67.5%; Pred. No. 2;
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 546 cctacgaaccacactgtgagtgcagagtcgacagaccttaaaaaaagaa 605
      |||||  |||  |||  |||||  |||  |||  |||  |||  |||  |||
Db 81 CCTACCACTCCCACTGTGCGTGGAGAATAATAGTCTTCTCCTCCCAAAAAAAAA 22

QY 606 aaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 21 AAAAAAAAAAAAAAAAAA 5

RESULT 11
M91492/c      219 bp      mRNA      EST      29-OCT-1992
LOCUS      HUMRTPEAK Homo sapiens CDNA.
DEFINITION      M91492
ACCESSION      M91492
NID      q275908
KEYWORDS      EST.
SOURCE      Human library-Subtracted human retinal pigment epithelium (RPE).
ORGANISM      Homo sapiens
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
1 (bases 1 to 219)
Expressed sequence tags and chromosomal localization of cDNA clones
from a subtracted retinal pigment epithelium library
Genomics 13 (3), 873-876 (1992)
92347897

JOURNAL
MEDLINE
COMMENT
Contact: Swaroop, A. and Giesler, L.
Department of Ophthalmology
Kellogg Eye Center
University of Michigan, Ann Arbor, MI 48105.
Location/Qualifiers
1..219
/organism="Homo sapiens"

BASE COUNT      36 a      52 c      54 g      77 t
ORIGIN

FEATURES
source
1..219

Query Match      5.9%: Score 36.8; DB 10; Length 219;
Best Local Similarity 63.6%; Pred. No. 2.1;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 535 cccacccctccctacgaaccacactgtgagtgcagagtcgacagaccttaaa 594
      |||||  |||  |||  |||||  |||||  |||  |||  |||  |||  |||
Db 97 CCGCCACACTGTGAGGTCCACACAGCCTATTAAAGGGAGATGTACTGCCAAAAAAA 38

QY 595 aaaaaaagaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 37 AAAAAAAAAAAAAAAAAAAAAA 10

RESULT 12
W37991/c      462 bp      mRNA      EST      15-MAY-1996
LOCUS      zc13609.s1 Soares parathyroid tumor NHPA Homo sapiens CDNA clone
DEFINITION      322216 3', mRNA sequence.
ACCESSION      W37991
NID      g1319604
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

REFERENCE 1 (bases 1 to 462)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
Mashu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 322.
Location/Qualifiers
1..462
/organism="Homo sapiens"
/note="vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5'-.
TGTACCAATCTGAAAGTGGAGCGCGCACACATTTTTTTTTTTTTTTTTTTT
T-3']", double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fátima Bonaldo. RNA from sporadic parathyroid adenomas was kindly provided by Dr. Stephen Marx, National Institute of Diabetes and Digestive and Kidney Diseases, NIH."
/db_xref="taxon:9606"
/clone="322216"
/clone_lib="Soares parathyroid tumor NbHPA"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
complement(1..>462)
128 a 97 c 109 g 125 t 3 others

BASE COUNT
ORIGIN
mRNA

Query Match 5.9%; Score 36.8; DB 13; Length 462;
Best Local Similarity 85.4%; Pred. NO. 2.3;
Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 575 aagtcgacgacgaccttaaaaaaaaaaagaaaaaacacacaaaa 622
|||
Db 48 AATGCGATCGACACTTTAAAAA
AAAAAAAAAAAAAAAAAAAAA 1

RESULT 13
LOCUS AA643056 283 bp mRNA EST 27-OCT-1997
DEFINITION nr5603.61 NCI-CGAP_P725 Homo sapiens cDNA clone IMAGE:1175740.
AA643056
ACCESSION
NID 92568274
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 283)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,
Ph.D.

CDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LMW at:
www-bio.liml.gov/dbfp/image/image.html

Seq primer: -40ml3 fwd, ET from Amer sham
High quality sequence stop: 273.
Location/Qualifiers

SOURCE

1..283
organism="Homo sapiens"
/note="Organ: prostate; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Normal prostate epithelial cell line (HPV
immortalized). 5' adaptor sequence: 5' GAATTCGCACAG 3'
3' adaptor sequence: 5' CTGACGTTTTTTTTTTTTTTT 3'
Average insert size: 1.1 kb."
/db_xref="taxon:9606"
/clone_image="IMAGE:1175740"
/clone_lib="NCI CGAP PR25"
/tissue_type="epithelium (cell line)"
/lab_host="Stor (kanamycin resistant)"

BASE COUNT

60 a 62 c 71 g 90 t

ORIGIN

Query Match 5.9%; Score 36.8; DB 24; Length 283;
Best Local Similarity 63.6%; Pct. No. 2.2;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Oy 535 ccacccctcccttcgcgaccacactgtgatgcatgcccagcgacaccttaa 594
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100 CCTCCACTCTCGGAGGTGCCACACGGCTATTAAAGGGCAATTCTACTCAGAAAAAAA 41
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40 AAAAAAAAAAAAAAAAAAAAAAAAAA 13

Oy 595 aaaaaaaagaaaaaaaaaacacaacaaa 622
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40 AAAAAAAAAAAAAAAAAAAAAAAAAA 13

RESULT 14

AA599770/c 412 bp mRNA EST 02-MAR-1998
LOCUS ag33all.s1 Jia bone marrow stroma Homo sapiens cDNA clone 1091324
DEFINITION 3' mRNA sequence.
ACCESSION AA599770
NID 92433395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 412)
Jia,L., Wilkin,D., Blitner,M., Robey,P., Young,M., Yamada,Y.,
Krizman,D., Liotta,L., Bonner,R., Schuler,G., Boguski,M.,
Powell,J., Lennon,G., Goodman,D., Hotchkiss,R., Melter,P.,
Trent,J., Hallier,L., Allen,M., Bowles,L., Geisler,S., Kucaba,T.,
Matta,M., Martin,J., Stepien,M., Tan,F., Theising,B., Bowers,Y.,
Wyllie,T., Waterston,R., Wilson,R. and Francomano,C.
WashU-MGB/NHGRI EST Project
Unpublished (1997)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK / Jia L
WashU-MGB/NHGRI EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1927 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amerisham
 High quality sequence stop: 362.
 Location/Qualifiers

FEATURES

source

```

1. 412
/organism="Homo sapiens"
/Note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI;
mRNA made from human bone marrow stroma, cDNA made by
oligo-dt priming. Directionally cloned. Size-selected for
average insert size >0.5 kb. Library supplied by Dr. Libin
Jia (NHGRI)."
/db_xref="taxon:9606"
/clone_1lb="Jia bone marrow stroma"
/clone_1lb="1091324"
/sex="mixed"
/tissue_type="bone marrow stroma"
/seq_stage="mixed"
/lab_host="XLI-Blue MRF"/SOER"
BASE COUNT      122 a      83 c      86 g      121 t
ORIGIN

```

Query Match
 Best Local Similarity 85.4%; Pred. No. 2.3; Length 412;
 Matches 41; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 575 aagtcgacgacgaccttaaaaaaacaacacaaaa 622
 ||||||| ||| ||||||||||| ||||||| ||| |||||
 Db 51 AATGCAATCAGCAGCTTAAAAAATAAAAAATAAAAA 4

RESULT 15
 AA804614 337 bp mRNA EST 19-FEB-1998
 LOCUS ob98b04.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1339375,
 DEFINITION mRNA sequence.
 ACCESION AA804614
 NID 92876015
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 337)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.lnl.gov/db/rp/image/image.html

Insert Length: 879 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amerisham
 High quality sequence stop: 289.

FEATURES

source

```

1. 337
/organism="Homo sapiens"
/Note="Vector: pT773D-Pac (Pharmacia) with a modified

```

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-GTGACCAATCTGAAGTGGGAGGCGCCGCTCATTTT-3']
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT773 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /clone_image:1339375
 /clone_1lb="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 BASE COUNT 71 a 69 c 91 g 106 t
ORIGIN

Query Match
 Best Local Similarity 63.6%; Pred. No. 2.3;
 Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 535 cccacctctccctacgaaccacactgtgagtgcacagtcgacgacaccttaa 594
 ||| ||||| | ||||| | | | | | | | | | | |
 Db 114 CCTCCACCTCTGGAGTCCACACAGCTATTAAAGGAAATGTACTCAAAAAA 55
 Qy 595 aaaaaaagaaaaaaacacacaaaaa 622
 ||||||| ||||||| ||| |||||
 Db 54 AAAAAAAAAAAAAAAAAAAAAAAAAA 27

Search completed: May 13, 1999, 19:44:17
 Job time: 11322 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:35:05 ; Search time 2254.51 Seconds

(without alignments)
1999.419 Million cell updates/sec

Title: US-09-212-270-10

Perfect score: 1260

Sequence: 1 ctctctagaggagctgcgaac.....tgcgtgtgtgtatagatgagg 1260

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database :

GenEmbl: *
1: gb_dal: *
2: gb_dal: *
3: gb_dal: *
4: gb_dal: *
5: gb_dal: *
6: gb_dal: *
7: gb_dal: *
8: gb_dal: *
9: gb_dal: *
10: gb_dal: *
11: gb_dal: *
12: gb_dal: *
13: gb_dal: *
14: gb_dal: *
15: gb_dal: *
16: gb_dal: *
17: gb_dal: *
18: gb_dal: *
19: gb_dal: *
20: gb_dal: *
21: gb_dal: *
22: gb_dal: *
23: gb_dal: *
24: gb_dal: *
25: gb_dal: *
26: gb_dal: *
27: gb_dal: *
28: gb_dal: *
29: gb_dal: *
30: gb_dal: *
31: gb_dal: *
32: gb_dal: *
33: gb_dal: *
34: gb_dal: *
35: gb_dal: *
36: gb_dal: *
37: gb_dal: *
38: gb_dal: *
39: gb_dal: *
40: gb_dal: *
41: gb_dal: *
42: gb_dal: *
43: gb_dal: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query
No. Score Match length DB ID Description

1	1233.2	97.9	1348	42	AF046888	AF046888 Homo sapi
2	1233.2	97.9	1348	42	AF046888	AF046888 Homo sapi
3	300	23.8	405	43	G37120	G37120 SHGC-56910
4	110	8.7	288	40	HS40D10F	260980 H.sapiens C
5	110	8.7	288	40	HS40D10F	260980 H.sapiens C
6	68.2	5.4	7218	6	166494	166494 Sequence 14
7	44.6	3.5	7218	6	166494	166494 Sequence 14
8	44.4	3.5	223	10	HS40D10R	260981 H.sapiens C
9	44.4	3.5	223	10	HS40D10R	260981 H.sapiens C
10	42.4	3.4	256404	18	AC005071	AC005071 *** SEQUE
11	41.6	3.3	116917	12	AC003024	AC003024 Human Chr
12	41.6	3.3	116917	42	AC003024	AC003024 Human Chr
13	41.2	3.3	33233	10	HSU51280	U51280 Human chrom
14	41.2	3.3	53522	11	HUMPRK1GEN	L39891 Homo sapien
15	41.2	3.3	200906	18	HSU51280	U51280 Human chrom
16	41.2	3.3	33233	40	HSU51280	U51280 Human chrom
17	41.2	3.3	53522	41	HUMPRK1GEN	L39891 Homo sapien
18	41.2	3.3	111863	18	HS73E16	295330 Human DNA s
19	40.6	3.2	37005	11	AC005239	AC005239 Homo sapi
20	40.6	3.2	135046	11	HSU91325	U91325 Human chrom
21	40.6	3.2	37005	41	AC005239	AC005239 Homo sapi
22	40.6	3.2	135046	41	HSU91325	U91325 Human chrom
23	40	3.2	317311	18	AC004085	AC004085 *** SEQUE
24	39.8	3.2	4433	40	HUMINCPSS	M19171 Human cyste
25	39.8	3.2	4433	40	HUMINCPSS	M19171 Human cyste
26	39.6	3.1	57493	11	AC004505	AC004505 Homo sapi
27	39.6	3.1	202004	11	HUAF001549	AF001549 Human Chr
28	39.6	3.1	57493	41	AC004505	AC004505 Homo sapi
29	39.6	3.1	202004	41	HUAF001549	AF001549 Human Chr
30	39.4	3.1	47991	12	AC005775	AC005775 Homo sapi
31	39.4	3.1	112444	18	AC003022	AC003022 *** SEQUE
32	39.4	3.1	47991	42	AC005775	AC005775 Homo sapi
33	39.2	3.1	102703	11	AC004910	AC004910 Homo sapi
34	39.2	3.1	112846	11	AC004910	AC004910 Homo sapi
35	39.2	3.1	151630	11	AC005154	AC005154 Homo sapi
36	39.2	3.1	151630	11	AC005154	AC005154 Homo sapi
37	39.2	3.1	102703	41	AC004910	AC004910 Homo sapi
38	39.2	3.1	112846	41	AC004910	AC004910 Homo sapi
39	39.2	3.1	151630	41	AC005154	AC005154 Homo sapi
40	39.2	3.1	151630	41	AC005154	AC005154 Homo sapi
41	39	3.1	20521	11	HSMECNH2	246773 H.sapiens D
42	39	3.1	177236	18	AC004581	AC004581 *** SEQUE
43	39	3.1	177236	18	AC005143	AC005143 *** SEQUE
44	39	3.1	20521	41	HSMECNH2	246773 H.sapiens D
45	38.8	3.1	101072	42	AC005752	AC005752 Homo sapi

ALIGNMENTS

RESULT	1	AF046888	1348 bp	mrna	PRI	26-SEP-1998
LOCUS	AF046888					
DEFINITION		Homo sapiens proliferation inducing ligand APRIL mRNA, complete cds.				
ACCESSION	AF046888					
NID	93650491					
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS		Hahne,M., Kataoka,T., Schreuter,M., Hofmann,K., Irmeler,M., Boder,J., L., Schneider,P., Bonnard,T., Holler,N., French,L.E., Soriat,B., Rimoldi,D., and Tschopp,J.				
TITLE		APRIL, a new ligand of the tumor necrosis factor family, stimulates tumor cell growth				
JOURNAL		J. Exp. Med. 188 (6), 1185-1190 (1998)				
MEDLINE		98416181				
REFERENCE		2 (bases 1 to 1348)				
AUTHORS		Hahne,M. and Tschopp,J.				
TITLE		Direct Submission				

JOURNAL Submitted (04-FEB-1998) Biochemistry, University of Lausanne, Ch.
des Boveresses 155, Epalinges, VD 1066, Switzerland

FEATURES
source
Location/Qualifiers
1. 1348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="uterus"
282. 1034
/note="TNF ligand family"
/codon_start=1
/product="proliferation inducing ligand APRIL"
/db_xref="P01935"50492"
/translation="MPASSPFLAPKPPENMGPPREPLASVALMLWGALGAVAC
AAALVQTEELSLREVSRLQGTGSPQNGEGPMQSLDEQSDALEAENGSRK
RAVLLQKOKHSHLVLP INATSKDSQVTEYMQPALRRGGLDAQGYRIQDA
GYLLYSQVLEFQDFTMGQVVSREGGROETFRCLRSMPSHSDRAYNCSYAGVTH
LHGGDLSVILIPRAKLNLSPHOTFEGFVPL"
misc_feature
366. 428
/note="transmembrane-region site"

BASE COUNT 293 a 377 c 374 g 304 t
ORIGIN

Query Match 97.98; Score 1233.2; DB 12; Length 1348;
Best Local Similarity 99.68; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

```

QY 1 ctctctagaaggacttggagacttaattctcctgtagctgaggagaggtgaggtctcaag 60
Db 10 CTCTCTAGAGGAGATGTGAACTTAATTCCTGAGAGCTGAGGAGGAGGAGGCTCTCAAG 69
QY 61 gcaacgctgagcccaagcagagtgctccagagacctaacaagtaacctagctgtctcc 120
Db 70 GCAACGCTGAGCCCAAGCAGAGTGCCAGAGCACTAACACTACCTTACCTGTGCTTCC 129
QY 121 tctcctcctcttatttatttcaagttctctttatttctccttgctgaacaacctcttc 180
Db 130 TCTCTCCTCTTTTATTTATTAAGTTCCTTTTATTTCTCTGCGTAACAACCTTCTTC 189
QY 181 cctctgacacactgcccgtacaccttaacggcccgcacactctctgtataccaccttt 240
Db 190 CCTTGTGACACACTGCGGTACCTTACCTTACCGCGCCCGCACTCTGCTGACCCACTCTT 249
QY 241 gaacacacagctgttggcagaggtcccaagctcatgacagctcatctctcttctgtag 300
Db 250 GAACACACAGCTGTGGAGAGGTCCCAAGCTCATGCGCCTCATCTCTTCTTCTGCTAG 309
QY 301 ccccccaga-ggctccagagcaacatgggggcccagltcagagagccggactctcaagttg 359
Db 310 CCCCCAAGGCGCTCCAGGCAACATGGGGGCCCAAGTCAGAGACCGGCACTCTCAGTTG 369
QY 360 cccctggttgaattgggggagcagctcctgtagggcggtgtgtg-catgtttcaagctga 418
Db 370 CCTCTGTGTTAGTTGGGGGAGCTCTGGGGGCGGTGGCTTGGCCATGGCTCTGCTGA 429
QY 419 cccaacaacaagactcagagacctcagagagagaggtgagccgctgtagagggagagag 478
Db 430 CCCAACAACAAGACTGAGAGCCTCAGAGAGAGAGTGAAGCGCGCTGAGGGAGACAGAG 489
QY 479 gccctcccaagaatggggaaggtatccctgtagcagagctcctccgagcagagttccgagt 538
Db 490 GCCCTCCCAAGATGGGGAAGGTATCCCTGAGAGTCTCCGAGAGAGAGTTCGAGTG 549
QY 539 cccctggaagcctgtagagagatggagagagatcccgaaaagagagcagtgctcacccaaa 598
Db 550 CCCCTGGAAGCCTGGAGAGATGGGAGAGATCCGGAAAAAGAGCAATGCTCACCCAAA 609
QY 599 aacagaagaagcagacactctgctcagcactggttcccatlaagccacccccaagagtg 658
Db 610 AACAGAAGAAGCAGACACTCTCTCTGACACTGTGTTCCATTAAAGCCCACTCCAAAGATG 669
QY 659 actccgattgtacagaggtgattgtgcaacacagctcttaaggcgtgtagagagcctacag 718

```

Db 670 ACTCCGATGTGACAGAGTGATGTGCAACACGCTTTAGCGTGGGAGAGCCCTACAGG 729
QY 719 cccaagatatgtgttcggaatccagagatgctggaattatctgctctatagccaggtcc 778
Db 730 CCCAAGATATGTGTGCCAATCCAGAGATGCTGAGATTATCTGCTATATAGCCAGCTCC 789
QY 779 tgttcaagacgttacttcaaccatgggtcaggtggtgtctcgaagagggccaagagagc 838
Db 790 TGTTCAAAGAGTACTCTTACCACTGGGTGAGGTGTCTCTCAGAAAGGCCAAGAAAGGC 849
QY 839 aggaagactctattccagatgataaagaatgacctcccccaccccgagccgtacacaa 898
Db 850 AGGAGACTCTATTCCAGATATAAAGATATGCCCTCCACCCGAGCCGAGCTTACACA 909
QY 899 gctcctatagcagaggtgttccatttaacccaagggatattctgagtgataattc 958
Db 910 GCTGCTATAGCGCAGGTGTCTCCATTACACCAAGGGGATATTCAGTCAATTAATTC 969
QY 959 cccggcgaagggcnaaacttaacctctctccacatggaacctctcgtgggttgtagaac 1018
Db 970 CCGGCGCAAGGGCAAACTTAACTCTCTCCACATGGAACCTTCTGGGTTGTGAAC 1029
QY 1019 tgtgattgtgtataaaagtggtcccaagcttggaaagccaggttgatatactgta 1078
Db 1030 TGTGATTGTGTATAAAAGTGCGTCCAGCTTGAAGACCAAGGTGGTATACATCTGA 1089
QY 1079 gacagccaagagctgagatataataaagagagagatgtgaaagaaagagcatcttct 1138
Db 1090 GACAGCCAAGAGCTGATATATAAAGAGAGGAAATGTGAGAAACAGAGCACTTCTCT 1149
QY 1139 gggttgtctcccgctcccaacttcccttcttcaattcccaacccttagacttgatt 1198
Db 1150 GGGTTGTGCTCCCGTTCCTCATCTTCCCTTTTCATCTCCACCCCTTAAGCTTGAATT 1209
QY 1199 taaggatacttgctctgttccccaatgagagctccgaattcttgctgtgtatagatga 1258
Db 1210 TACGATATCTTGTCTGTCCCATGAGCTCCGAATCTTGTGCTGTGATGATGAG 1269
QY 1259 gg 1260
Db 1270 GG 1271

RESULT 2
AF046888
LOCUS AF046888 1348 bp mRNA PRI 26-SEP-1998
DEFINITION Homo sapiens Proliferation inducing ligand APRIL mRNA, complete cds.
ACCESSION AF046888
NID 93650491
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 1348)
Hahne, M., Kataoka, T., Schroeter, M., Hofmann, K., Irmeler, M.,
Bodmer, J., U., Schneider, P., Bornand, T., Holler, N., French, L.E.,
Sordal, B., Rimoldi, D., and Tschopp, J.
APRIL, a new ligand of the tumor necrosis factor family, stimulates
tumor cell growth
J. Exp. Med. 188 (6), 1185-1190 (1998)
JOURNAL
MEDLINE 98416181
REFERENCE 2 (bases 1 to 1348)
Hahne, M. and Tschopp, J.
AUTHORS Direct Submission
TITLE Submitted (04-FEB-1998) Biochemistry, University of Lausanne, Ch.
JOURNAL des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
source
Location/Qualifiers
1. 1348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="uterus"


```

CDs      282..1034
         /note="TNF ligand family"
         /codon_start=1
         /product="proliferation inducing ligand APRIL"
         /db_xref="PID:3650493"
         /translation="MPASSPFLAPKPGPPNGGPPREPALVALMLSMGALGAVAC
         AMALTOOTERLOSLREVSRLOGTGGPSONGNGEYMOSEPPSSDALEMEWGESRK
         RBAVLTOKOKHOSVILHYLPYRNATSKDSVDYEVNKKPRLRGRLQAGVYRIODA
         GYLLISQVLFODVITFMQVVSREGQOETLFRICRSPSHPDPAVNSCSAGVFH
         LHQDILSVILIRARAKMLSPHGFLEFVKL"
         366..428
         /note="transmembrane-region site"
BASE COUNT 293 a 377 c 374 g 304 t
ORIGIN
Query Match      97.9%; Score 1233.2; DB 42; Length 1348;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY      1 ctctcagagagagctggaactaattctcctcagagctgagagagctgagagctcgaag 60
        |||||||
DB      10 CTTCCTAGAGGGGACTGGAACCTAATCTCTCTGAGCTGAGGAGGAGGCTGAGGCTCTCAAG 69
        |||||||

QY      61 gcaagcctgagcccaagagagctgcaagagagactaaagtaaccttgccttgccttcc 120
        |||||||
DB      70 GCAACGCTGGCCCCACAGAGGAGTGCAGAGACATACAGTACCTTACCTTGCCTTTC 129
        |||||||

QY      121 tctcctcctcttcttcttctcagctccttcttcttctcctcctcctcctcctcctcctc 180
        |||||||
DB      130 TCTCTCCTCTCTTTTATTTTCAAGTCTCTTTTATTTCTCTTCTGCTGCTACACTTCTTC 189
        |||||||

QY      181 cctctcagacactgcgcgtacacctaccgcgcgcgcacactcctcctcctcctcctcct 240
        |||||||
DB      190 CCTTCTGCACACACTGCCGATACCTTACCCTGCGCCGACCTCTCTTCTGCTACCTTCTT 249
        |||||||

QY      241 gaaacacagcgtgttgagcaggggtcccaagctctatgcacagctctatctccttctctag 300
        |||||||
DB      250 GAAACACAGCTGTGTGCGAGGGTCCACAGCTCATGCTACAGCTCATCTTCTTCTCTAG 309
        |||||||

QY      301 ccccccaga-ggcctccagacaacatgagggggccagctcagagagccggcactcctcagttg 359
        |||||||
DB      310 CCCCAGAAAGGCTCCAGGACATAGGGGGCCCACTCAGAGAGCCGGCCTCTCAGTTG 369
        |||||||

QY      360 cccctcgtgtgagctgaggggagcctcctcggggcgtggtgtg-catggttcagctga 418
        |||||||
DB      370 CCTCTGCTGTGAGTTGGGGGCGAGCTCTGGGGCGCGTGTGTCATGCGTCTGCTGTGA 429
        |||||||

QY      419 ccccaaaaacagagctgcagagagctcagagagagaggtgagccgctgcagggagagagag 478
        |||||||
DB      430 CCCAACAAACAGAGCTGCAGAGCTCAGGAGAGAGGTGAGCCCGCTGCAGGGGACAGGAG 489
        |||||||

QY      479 gcccccacagaatgggagaggtatccctcgtgagagctcccgaggagcaaggtccagatg 538
        |||||||
DB      490 GCCCCCTCCAGATAGGGAGAGGTATCCCTGGAGAGACTCTCCGGAGACAGAGTCTCCGATG 549
        |||||||

QY      539 cccctggaagccttggagagatgggagagagatcccgaaagagagagagctgctaccacaa 598
        |||||||
DB      550 CCTTGGAAGCCTGGAGATGGGAGAGATCCCGAAAGAGAGACAGTCTCTACCCAAA 609
        |||||||

QY      599 aaaaagaagaagcagcactgtctcctcgaacctggttcccatlaagcccaactccaagatg 658
        |||||||
DB      610 AACAGAAGAGCAGCACTGTCTCTCACCCTGTTCCCATTAACGCCACCTCCAAAGATG 669
        |||||||

QY      659 actcgaatgtacagaggtatgttgcaacagctcttaagcgttggagagagagctcagag 718
        |||||||
DB      670 ACTCCCATGTGACAGAGGTATGTGCGAACCACTCTTAAGCGCTGGAGAGGCTTACAGG 729
        |||||||

QY      719 cccaagagataggctgctcgaatccagagatgctgagattatctgctgtatagccaagctc 778
        |||||||
DB      730 CCCAAGGATATGCTGCCGAATCCAGAGATCTGAGATTATCTGCTGTATAGCAGAGTCC 789
        |||||||

QY      779 tgtttaagaagcgtgactttcaacatggtgtgtgtctcgaagaagccaaagaagc 838

```

```

DB      790 TGTTTAAGACGTGACTTTCCACCATGCTGCTAGCTGTCTCTCAGAAAGCCAAAGAGGC 849
        |||||||
QY      839 aggaagctatctccagatgtataagaagatgcctcccaaccggagccggctacaaaca 898
        |||||||
DB      850 AGGAGACTCTATTCGATGTATAGAGATATGCCCTCCACCCGGAGCCGGCTCTACACA 909
        |||||||

QY      899 gctcctatagcagagtgctctccatcttaacacaaaggagatattcgtagtgctaaatc 958
        |||||||
DB      910 GCTGTATAGCGCAGGTGCTTCTCATTTACACCAAGGAGATATTCTGAGTGTCTATATTC 969
        |||||||

QY      959 cccgggcaaggcggaactaactctctccacatgaaacttcctcgggtgtgtgaaac 1018
        |||||||
DB      970 CCGGGCAAGGGCGAAGACTTAACCTCTCTCCACATGGAACCTTCTGGGGTGTGTAAC 1029
        |||||||

QY      1019 tgtgattgttataaaagtgcctccagcttgaagaaccaagggtgtgtacatactga 1078
        |||||||
DB      1030 TGTGATTGTGTTATATAAAGTGGCTCCAGCTTGAAGACAGAGTGGGTACTACTGGA 1089
        |||||||

QY      1079 gaaagcaagagctgagatataaaggagagagaaatgtcaggaagagaggaacttctcct 1138
        |||||||
DB      1090 GACAGCCAAGAGCTGAGTATATTAAGAGAGGGAATGTGCAGGACAGAGGCACTTCTCC 1149
        |||||||

QY      1139 ggtgtgctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1198
        |||||||
DB      1150 GGTGTGGCTCCCGTCCCTCACTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTT 1209
        |||||||

QY      1199 taagatattctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1258
        |||||||
DB      1210 TAGGATATCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1269
        |||||||

QY      1259 gg 1260
        ||
DB      1270 GG 1271

RESULT      3
G37120/c      405 bp      DNA      STS      30-MAR-1998
LOCUS      SHGC-56910 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION      G37120
ACCESSION      G2996771
NID      92996771
KEYWORDS      STS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 405)
Myers,R.M.
Human STS (1997)
Unpublished (1997)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@hgc.stanford.edu
Primer A: ACACACACGCAAGAAATTCG
Primer B: AGGAGAGGAGATGTCAAGA
STS size: 149
PCR Profile:
Initial Incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:
Template: 25 ng
Primer: each 1 uM

```

dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer: MgCl₂: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs derived from AA057069 -- Unigene.

FEATURES
 source
 Location/Qualifiers
 1..405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="17"
 /clone_lib="Human"

SNS
 primer_bind
 primer_bind complement(195..214)
 BASE COUNT 107 a 110 c 99 g 88 t 1 others
 ORIGIN

Query Match 23.8%; Score 300; DB 43; Length 405;
 Best Local Similarity 96.8%; Pred. No. 2e-77;
 Matches 338; Conservative 0; Mismatches 6; Indels 5; Gaps 3;

QY 917 tcttcattacacacagggagatattctgagtg-cataattcccgaggcaaggc--ga 973
 |||
 Db 405 TCTTCATTACACACAGGAGATATTCTGAGTGCATTAATCCCGGCAAGGCGCGA 346
 QY 974 aactaacctctctcc--aacatgaacctctctgggttgtagaactgtgattgttta 1031
 |||
 Db 345 AACTTAACCTCTCTCCACATGSAACCTTCTGCGTTGTGTAACCTGTGATTGTGTTA 286
 QY 1032 taataagtggtcccaagcttggaaagaccagggtgtgtacatactggaagaccagaagc 1091
 |||
 Db 285 TAAAGATGGCTCCACACTTGGAGAGACAGGGTGGTACATACAGAGACCAAGAGC 226
 QY 1092 tggatataaaggagaggaatgtgcaggaacagagacatctctgggttggtctcc 1151
 |||
 Db 225 TGAATATATAAAGAGAGGAGATGTGACAGAGACAGCGCTCTCTGCGTTGGCTCC 166
 QY 1152 cgttcctcacttttccttttcattccacacccctagactttgatttaagagatcttg 1211
 |||
 Db 165 CGTTCCTCACTTTTCCCTTTTCATTCACCCCTTAAGCTTTGATTTACGAGATATCTTG 106
 QY 1212 ctctgttcccatgagagctccgaattcttgcgtgtgtgtagatgaagg 1260
 |||
 Db 105 CTCTGTCTCCCATGAGAGCTCCGAATCTTGCGTGTGTAGATGAGAGG 57

RESULT 4
 LOCUS HS40D10F 288 bp DNA PRI 22-OCT-1995
 DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10,
 forward read cp940d10.ftla.
 ACCESSION 260980
 MID g1033358
 KEYWORDS Cpg island; genomic MseI fragment.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Eukaryotes; Eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 288)
 AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
 REFERENCE 2 (bases 1 to 288)
 AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
 TITLE Purification of cpg islands using a methylated DNA binding column

JOURNAL Nature Genet. 6 (3), 236-244 (1994)
 MEDLINE 94282070
 COMMENT Vector: pGEM-5zf(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
 http://www.hgmp.mrc.ac.uk/ for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source
 Location/Qualifiers
 1..288
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /sex="male"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /clone="40d10"

BASE COUNT 77 a 68 c 85 g 57 t 1 others
 ORIGIN

Query Match 8.7%; Score 110; DB 10; Length 288;
 Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gatgactcagatgtgacagagtgatgtgcaaccagctcttagtgaggaagcccta 714
 |||
 Db 142 GATGACTCCGATGTGACAGAGTGATGTGCAACGACGCTTATGGCGTGGAGAGCCCTA 201
 QY 715 caggcccaagagatagtgtccgaatccagagatgctgagattatctgtc 764
 |||
 Db 202 CAGGCCCAAGATATGATGTGCCAATCCAGAGATGCGAGTTATCTGCG 251

RESULT 5
 LOCUS HS40D10F 288 bp DNA PRI 22-OCT-1995
 DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10,
 forward read cp940d10.ftla.
 ACCESSION 260980
 MID g1033358
 KEYWORDS Cpg island; genomic MseI fragment.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Eukaryotes; Eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 288)
 AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
 REFERENCE 2 (bases 1 to 288)
 AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
 TITLE Purification of Cpg islands using a methylated DNA binding column
 JOURNAL Nature Genet. 6 (3), 236-244 (1994)
 MEDLINE 94282070
 COMMENT Vector: pGEM-5zf(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
 http://www.hgmp.mrc.ac.uk/ for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source
 Location/Qualifiers
 1..288
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /sex="male"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /clone="40d10"

BASE COUNT 77 a 68 c 85 g 57 t 1 others
 ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gatgactccgagtgtgacagagtgatgtgcaacacgctcttaagcgttggagagccta 714
142 GATGACTCCGATGTGACAGAGTGATGTGCAACACGCTCTTAAGCGTGGAGAGGCCTA 201

QY 715 caggcccaagagatggtgttcggaatccagagatgcttgagtcttctgct 764
202 CAGGCCCAAGAGATGGTGTTCGGAATCCAGAGATGCTGAGTTATCTGCT 251

RESULT 6
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 7218)
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1. 7218
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others

Query Match
Best Local Similarity 11.8%; Pred. No. 2.5e-09;
Matches 34; Conservative 155; Mismatches 98; Indels 0; Gaps 0;

QY 19 acctaatctctcgtagagctgaggaggtgaggtctcaaggcaacgctggcccaaga 78
991 AACATATTTTCTCTGTTGCGCATACGCTCACAGAAATTAATTCGAGCTTGCTGCAGGT 1050

QY 79 cggagtgccgaggaacactaacagtaacctgtgcttctctctctctctcttatt 138
1051 CGAGGAGCTTGCATATTT 1110

QY 139 ttcagttcctcttattctctctcctgcgtacacactctctctctctgcacactgcc 198
1111 YY 1170

QY 199 gtaccttaccgcccgcacacactcctctgtacacacactcttgaaacacacagcttggc 258
1171 YY 1230

QY 259 agggctcccgactcagtcacgacctcactctctctctctctgtagccccc 305
1231 YY 1277

RESULT 7
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 7218)
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;

FEATURES
source Location/Qualifiers
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match
Best Local Similarity 3.5%; Score 44.6; DB 6; Length 7218;
Matches 8; Conservative 216; Mismatches 155; Indels 0; Gaps 0;

QY 363 tctgttgatgttgggggagcctctggggccggttggtgacgttgcagtcgaccca 422
1442 TTTCGTACRR 1383

QY 423 acaacagagctgcagagcctcagagagagtgtagcgcgctgcagggagacagagccc 482
1382 RRR 1323

QY 483 ctcccaaatggggaaggtatcccttgccagagctcccgagagcagagttcgcattgcc 542
1322 RRR 1263

QY 543 ggaagcctggagagatgggagagatcccgaaagagagcagtgctcaccacaaca 602
1262 RRR 1203

QY 603 gaagaagcagcactctctcctgcacctggttccataacgccaccctccagagtagctc 662
1202 RRR 1143

QY 663 cgaatgtacagagtgatgtgtgcaacagcctcctagcgtggagagagcctacagccca 722
1142 RRR 1083

QY 723 aggataagtgctcgcaatc 741
1082 RRR 1064

RESULT 8
LOCUS HS40D10R/c 223 bp DNA PRI 22-OCT-1995
DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10, reverse read cpg40d10.rta.
ACCESSION 260981
NID 91033359
KEYWORDS Cpg island; genomic MseI fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Carnivora; Homnidae; Homo.

REFERENCE
1 (bases 1 to 223)
AUTHORS MacDonald, M., Huckle, E., Wilkinson, P. and Micklem, G.
TITLE Direct Submision
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail contact: humgeny@sanger.ac.uk

REFERENCE
2 (bases 1 to 223)
AUTHORS Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
TITLE Purification of cpg islands using a methylated DNA binding column
JOURNAL Nature Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT
Vector: pGEM-5Zf(-)
Clones are available from the UK MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL: <http://www.hgmp.mrc.ac.uk/> for details
or contact: biohelp@hgmp.mrc.ac.uk.
Location/Qualifiers
1. 223
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="adult"
/sex="male"

FEATURES
source

```

*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. It consists of 35 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
* as runs of N as a convenience only. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
*
1
1600: contig of 1600 bp in length
1618: gap of unknown length
1619
3213: contig of 1595 bp in length
3214
3231: gap of unknown length
3232
4831: contig of 1600 bp in length
4832
4849: gap of unknown length
4850
6665: contig of 1816 bp in length
6666
6683: gap of unknown length
6684
8472: contig of 1789 bp in length
8473
8490: gap of unknown length
8491
10001: contig of 1511 bp in length
10002
10019: gap of unknown length
10020
11650: contig of 1631 bp in length
11651
11668: gap of unknown length
11669
13497: contig of 1829 bp in length
13498
13515: gap of unknown length
13516
15944: contig of 2429 bp in length
15945
15962: gap of unknown length
15963
17496: contig of 1534 bp in length
17497
17515
19092: contig of 1578 bp in length
17515
19093
19110: gap of unknown length
19111
21089: contig of 1979 bp in length
21090
21107: gap of unknown length
21108
22787: contig of 1680 bp in length
22788
22805: gap of unknown length
22806
24419: contig of 1614 bp in length
24420
24437: gap of unknown length
24438
25932: contig of 1495 bp in length
25933
25950: gap of unknown length
25951
27541: contig of 1591 bp in length
27542
27559: gap of unknown length
27560
29831: contig of 2272 bp in length
29832
29849: gap of unknown length
29850
31952: contig of 2103 bp in length
31953
31970: gap of unknown length
33770: contig of 1800 bp in length
33771
33788: gap of unknown length
33789
35495: contig of 1707 bp in length
35496
35513: gap of unknown length
35514
37491: contig of 1978 bp in length

```

FEATURES	SOURCE	LOCATION/Qualifiers
BASE COUNT	64346 a 62984 c 62157 g 66233 t	684 others
ORIGIN	/db_xref="taxon:9606" /clone="RG161A02"	
FEATURES	Location/Qualifiers	
SOURCE	1..256404	
ORGANISM	/organism="Homo sapiens"	
LOCUS	AC003024	
DEFINITION	Human Chromosome 15q26.1 PAC clone pDU41616, complete sequence.	
ACCESSION	AC003024	
NID	93608160	
KEYWORDS	HTG.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 116917)	
TITLE	Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M., Buetler, J., Bumeister, R., Card, P., Desaillo, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, Y., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.	

JOURNAL REFERENCE AUTHORS	Unpublished 2 (bases 1 to 116917) Evans,G.A., Athanasiou,M., Bradbury,P., Brignac,S., Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
TITLE JOURNAL	Direct Submission Submitted (21-OCT-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE AUTHORS	3 (bases 1 to 116917) Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S., Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
TITLE JOURNAL	Direct Submission Submitted (17-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE AUTHORS	4 (bases 1 to 116917) Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Baether,J., Bumester,R., Card,P., desallboaf,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., Mcarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Slimson,S., Syed,M. and Ward,T.
TITLE JOURNAL	Submitted (15-SEP-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT	On Sep 16, 1998 this sequence version replaced g1:2695560. Further information regarding the map of this region or annotation of pjd41616 can be found at http://gestec.swmed.edu/chromos5.htm .
JOURNAL	IMPOrRANT: This submission contains the entire insert of clone pjd41616. pjd41616 comes from the RPCI-3 PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.56. In addition, attempts have been made to assure over 99% of consensus base calls consist of either double stranded coverage or 2 types of labeling chemistry on one strand.
FEATURES	CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 15q26.1 Bloom's syndrome/DNA polymerase gamma region. This region is mapped between STS D15S171 and D15S652 MARKER CONFIRMATION: STS sequence confirmed; WI-6813, D15S963, SHCC-9392
SOURCE	MAPPED CLONE OVERLAP: pjd422e7. Location/Qualifiers 1..116917 /organism="Homo sapiens" /db_xref="taxon:9606" 5376..5521 /rpt_family="MIR" complement(8545..8834) /rpt_family="Alu" 9913..10219 /rpt_family="MER2" 11974..12151 /rpt_family="MLT1" 15289..16374 /rpt_family="L1" 17666..17816 /rpt_family="L1" 18471..18642 /rpt_family="Alu" complement(20327..20615)

Repeat_region	/rpt_family="Alu"	complement(20947)	.21077)
Repeat_region	/rpt_family="MIR"	21210.	.21508
Repeat_region	/rpt_family="Alu"	23183.	.23483
Repeat_region	/rpt_family="Alu"	complement(26648)	.27241)
Repeat_region	/rpt_family="Alu"	complement(21278).	.27386)
Repeat_region	/rpt_family="MLT1"	30624.	.31046
Repeat_region	/rpt_family="MER42"	complement(32058)	.33119)
Repeat_region	/rpt_family="MER3"	complement(32746).	.33980)
Repeat_region	/rpt_family="MER43"	33389.	.37165
Repeat_region	/rpt_family="L1"	37567.	.38347
Repeat_region	/rpt_family="MER42"	complement(43991).	.44279)
Repeat_region	/rpt_family="Alu"	complement(44423).	.44623)
Repeat_region	/rpt_family="MIR"	complement(47160).	.47446)
Repeat_region	/rpt_family="Alu"	complement(49643).	.49937)
Repeat_region	/rpt_family="Alu"	50664.	.50947
Repeat_region	/rpt_family="Alu"	complement(52027).	.52266)
Repeat_region	/rpt_family="MER46"	complement(52470).	.52541)
Repeat_region	/rpt_family="MER46"	complement(58313).	.58613)
Repeat_region	/rpt_family="Alu"	59242.	.59382
Repeat_region	/rpt_family="MER46"	complement(59243).	.59474)
Repeat_region	/rpt_family="MER46"	61236.	.61507
Repeat_region	/rpt_family="Alu"	63584.	.63715
Repeat_region	/rpt_family="MERS"	complement(65384).	.63714)
Repeat_region	/rpt_family="MERS"	complement(70158)	.70391)
Repeat_region	/rpt_family="Alu"	72078.	.72148
Repeat_region	/rpt_family="L1"	72205.	.72466
Repeat_region	/rpt_family="Alu"	74340.	.74493
Repeat_region	/rpt_family="MIR"	complement(76577).	.76469)
Repeat_region	/rpt_family="MIR"	77426.	.77717
Repeat_region	/rpt_family="Alu"	complement(82461).	.82745)
Repeat_region	/rpt_family="Alu"	84074.	.84249
Repeat_region	/rpt_family="Alu"	complement(85332).	.85612)
Repeat_region	/rpt_family="Alu"	88044.	.88333
Repeat_region	/rpt_family="Alu"	90721.	.90757
Repeat_region	/rpt_family="MIR"	93074.	.93360
Repeat_region	/rpt_family="Alu"	complement(94986).	.95294)
Repeat_region	/rpt_family="Alu"		

repeat_region	98459.	.98706	
/rpt_family="L1"	complement(98873.	.99025)	
repeat_region	/rpt_family="MR"		
repeat_region	complement(98895.	.100145)	
repeat_region	/rpt_family="Alu"		
repeat_region	100174.	.100500	
repeat_region	/rpt_family="THE1"		
repeat_region	complement(100506.	.100539)	
repeat_region	/rpt_family="Alu"		
repeat_region	101454.	.101641	
repeat_region	/rpt_family="MLT1"		
repeat_region	complement(102058.	.102275)	
repeat_region	/rpt_family="Alu"		
repeat_region	106631.	.106929	
repeat_region	/rpt_family="Alu"		
repeat_region	complement(109874.	.111578)	
repeat_region	/rpt_family="L1"		
repeat_region	complement(112332.	.112482)	
repeat_region	/rpt_family="MER5"		
repeat_region	complement(112818.	.113102)	
repeat_region	/rpt_family="Alu"		
BASE COUNT	36124	a 24604	c 23503
ORIGIN		g 32686	t

[illegible]

TITLE	Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
JOURNAL	Direct Submission
REFERENCE	Submitted (21-Oct-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
AUTHORS	3 (bases 1 to 116917) Evans, G.A., Athanasiou, M., Basil, M., Bradbury, P., Brignac, S., Bumester, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Narayanaswamy, U., Newton, J., O'Brien, K., Patel, P., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
TITLE	Direct Submission
JOURNAL	Submitted (17-Dec-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE	4 (bases 1 to 116917)
AUTHORS	Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M., Buettner, J., Bumester, R., Card, P., desallboaf, F., Dunn, J., English, C., Ehrldge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
TITLE	Direct Submission
JOURNAL	Submitted (16-SEP-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT	On Sep 16, 1998 this sequence version replaced g1:2695560. Further information regarding the map of this region or annotation of pD41616 can be found at http://gestec.smed.edu/chromos05.htm . IMPORTANT: This submission contains the entire insert of clone pD41616. pD41616 comes from the Rpci-3 PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.56. In addition, attempts have been made to assure over 99% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 15q26.1 Bloom's syndrome/DNA polymerase gamma region. This region is mapped between STS D1S171 and D1S5652 MARKER CONFIRMATION: STS sequence confirmed; WI-6813, D1S5963, SHGC-9392
FEATURES	MAPPED CLONE OVERLAP: pD422e7. Location/Qualifiers
SOURCE	1.116917 /organism="Homo sapiens" /db_xref="taxon:9606" 5376..5521 /rpt_family="MIR" complement(8545..8834) repeat_region /rpt_family="Alu" 9913..10219 /rpt_family="MER2" 11974..12151 /rpt_family="MLT1" 15289..16374 repeat_region /rpt_family="L1" 17666..17816 repeat_region /rpt_family="L1" 18471..18642 repeat_region /rpt_family="Alu" complement(20327..20615) repeat_region /rpt_family="Alu" complement(20947..21077) repeat_region /rpt_family="MIR" 21210..21508 repeat_region /rpt_family="Alu" 23183..23483 repeat_region /rpt_family="Alu"
repeat_region	complement(26948..27241) /rpt_family="Alu"
repeat_region	complement(27278..27386) /rpt_family="MLT1"
repeat_region	30624..31046 /rpt_family="MER42"
repeat_region	complement(32058..32119) /rpt_family="MER43"
repeat_region	complement(32746..32980) /rpt_family="MER43"
repeat_region	33389..33765 /rpt_family="L1"
repeat_region	37567..38347 /rpt_family="MER42"
repeat_region	complement(43991..44279) /rpt_family="Alu"
repeat_region	complement(44423..44623) /rpt_family="MIR"
repeat_region	complement(47160..47446) /rpt_family="Alu"
repeat_region	complement(49643..49937) /rpt_family="Alu"
repeat_region	50664..50947 /rpt_family="Alu"
repeat_region	complement(52207..52266) /rpt_family="MER46"
repeat_region	complement(52470..52541) /rpt_family="MER46"
repeat_region	complement(58313..58613) /rpt_family="Alu"
repeat_region	59242..59382 /rpt_family="MER46"
repeat_region	complement(59243..59474) /rpt_family="MER46"
repeat_region	61236..61507 /rpt_family="Alu"
repeat_region	63584..63715 /rpt_family="MER5"
repeat_region	complement(63584..63714) /rpt_family="MER5"
repeat_region	complement(70158..70391) /rpt_family="Alu"
repeat_region	72076..72148 /rpt_family="L1"
repeat_region	72205..72466 /rpt_family="Alu"
repeat_region	74340..74493 /rpt_family="MIR"
repeat_region	complement(76377..76469) /rpt_family="MIR"
repeat_region	77426..77717 /rpt_family="Alu"
repeat_region	complement(82461..82745) /rpt_family="Alu"
repeat_region	84074..84249 /rpt_family="Alu"
repeat_region	complement(85332..85612) /rpt_family="Alu"
repeat_region	88044..88333 /rpt_family="Alu"
repeat_region	90721..90757 /rpt_family="MIR"
repeat_region	93074..93360 /rpt_family="Alu"
repeat_region	complement(94986..95294) /rpt_family="Alu"
repeat_region	98459..98706 /rpt_family="L1"
repeat_region	complement(98873..99025) /rpt_family="MIR"
repeat_region	complement(99895..100145) /rpt_family="Alu"
repeat_region	100174..100500

intron	/gene="PKD1" /number=3 19975. .20095 /gene="PKD1" /number=3 20096. .20167 /gene="PKD1" /number=4 20168. .20435 /gene="PKD1" /number=4 20436. .20605 /gene="PKD1" /number=5 20606. .21109 /gene="PKD1" /number=5 21110. .21490 /gene="PKD1" /note="putative" /number=6 21174. .21459 /gene="PKD1" /note="Cpg island; putative" 21491. .21608 /gene="PKD1" /number=6 21609. .21792 /gene="PKD1" /note="putative" /number=7 21793. .22227 /gene="PKD1" /number=7 22228. .22448 /gene="PKD1" /note="putative" /number=8 22449. .22636 /gene="PKD1" /number=8 22637. .22752 /gene="PKD1" /note="putative" /number=9 22753. .23117 /gene="PKD1" /number=9 23118. .23289 /gene="PKD1" /note="putative" /number=10 23290. .23655 /gene="PKD1" /number=10 23656. .23815 /gene="PKD1" /note="putative" /number=11 23816. .24966 /gene="PKD1" /number=11 24967. .25111 /gene="PKD1" /note="putative" /number=12 25112. .25988 /gene="PKD1" /number=12 25989. .26120 /gene="PKD1" /note="putative" /number=13	intron	26121. .26317 /gene="PKD1" /number=13 26318. .26493 /gene="PKD1" /note="putative" /number=14 26494. .26807 /gene="PKD1" /number=14 26808. .26941 /gene="PKD1" /note="putative" /number=15 26942. .27406 /gene="PKD1" /number=15 27407. .28254 /note="Cpg island; putative" 27407. .27525 /gene="PKD1" /note="putative" /number=16 27526. .27627 /gene="PKD1" /number=16 27628. .28683 /gene="PKD1" /note="putative" /number=17 28608. .28993 /note="Cpg island; putative" 28684. .29125 /gene="PKD1" /number=17 29126. .29875 /gene="PKD1" /note="putative" /number=18 29857. .30825 /note="Cpg island; putative" 29876. .30418 /gene="PKD1" /number=18 30419. .31024 /gene="PKD1" /note="putative" /number=19 31025. .31243 /gene="PKD1" /number=19 31244. .31393 /gene="PKD1" /note="putative" /number=20 31394. .32327 /gene="PKD1" /number=20 32328. .32471 /gene="PKD1" /note="putative" /number=21 32472. .32598 /gene="PKD1" /number=21 32599. .32878 /gene="PKD1" /note="putative" /number=22 32879. .33079 /note="Cpg island; putative" 32879. .32971 /gene="PKD1"
--------	---	--------	---

exon	/number=22 32972..33185 /gene="PKD1" /note="putative"	exon	/note="Cpg island; putative" 41523..41733 /gene="PKD1"
intron	/number=23 33186..33250 /gene="PKD1"	intron	/number=33 41734..41823 /gene="PKD1"
exon	/number=23 33251..33410 /gene="PKD1" /note="putative"	exon	/number=33 41824..41950 /gene="PKD1"
intron	/number=24 33411..33847 /gene="PKD1"	intron	/number=34 41951..43609 /gene="PKD1"
exon	/number=24 33848..33952 /gene="PKD1" /note="putative"	repeat_region	/number=34 42463..42685 /rpt_family="Alu"
intron	/number=25 33953..36953 /gene="PKD1"	repeat_region	/number=34 43084..43321 /rpt_family="Alu"
exon	/number=25 36954..37098 /gene="PKD1" /note="putative"	exon	/number=35 43610..43726 /gene="PKD1"
misc_feature	/number=26 37007..37251 /note="Cpg island; putative"	intron	/number=35 43727..43816 /gene="PKD1"
intron	/number=26 37099..37700 /gene="PKD1"	exon	/number=35 43817..43866 /gene="PKD1"
exon	/number=26 37701..38330 /gene="PKD1"	intron	/number=36 43867..44090 /gene="PKD1"
intron	/number=27 38331..38623 /gene="PKD1"	exon	/number=36 44091..44275 /gene="PKD1"
exon	/number=27 38624..38780 /gene="PKD1"	intron	/number=37 44276..44352 /gene="PKD1"
intron	/number=28 38781..38960 /gene="PKD1"	exon	/number=37 44353..44446 /gene="PKD1"
exon	/number=28 38961..39213 /gene="PKD1"	intron	/number=38 44447..47383 /gene="PKD1"
intron	/number=29 39214..39337 /gene="PKD1"	repeat_region	/number=38 44778..45002 /rpt_family="Alu"
exon	/number=29 39338..39533 /gene="PKD1"	repeat_region	/number=38 45003..46602 /rpt_family="Alu"
intron	/number=30 39534..41027 /gene="PKD1"	exon	/number=39 47384..47502 /gene="PKD1"
repeat_region	/number=30 complement(39970..40228) /rpt_family="Alu"	intron	/number=39 47503..47580 /gene="PKD1"
repeat_region	40311..40598 /rpt_family="Alu"	exon	/number=39 47581..47783 /gene="PKD1"
exon	41028..41198 /gene="PKD1"	intron	/number=40 47784..47855 /gene="PKD1"
intron	/number=31 41199..41284 /gene="PKD1"	exon	/number=40 47856..48050 /gene="PKD1"
exon	/number=31 41285..41428 /gene="PKD1"	intron	/number=41 48051..48500 /gene="PKD1"
intron	/number=32 41429..41522 /gene="PKD1"	exon	/number=41 48501..48640 /gene="PKD1"
misc_feature	/number=32 41501..41730	intron	/number=42 48641..49001 /gene="PKD1"
		exon	/number=42 49002..49114 /gene="PKD1"
		intron	/number=43 49115..49405

Feature	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
exon	49406.49547							
intron	49548.49687							
exon	49688.49813							
intron	49814.49996							
misc_feature	49898.50854							
exon	49997.50171							
intron	50172.50418							
exon	50419.50709							
conflict	50652.50653							
intron	50710.50784							
exon	50785.50919							
conflict	50796.50797							
intron	50920.51002							
exon	51003.51308							
intron	51309.51398							
exon	51399.52883							
conflict	51827.51828							
polya_signal	52862.52867							
BASE COUNT	8486 a 17664 c 15766 g 11606 t							
ORIGIN								

```
Db 35650 CTTCCTTTCCCTCCCTCTTCTCTTTCCTCTCCTCCCTTCTTTCCCTCTTCCCT 35909
QY 169 acaactcttcaccccttgcagaccagtccgcltacacctaacccgccgcacatccttgc 228
Db 35910 CCCCTCGCTGCCCCCTGCCCTCGCTGCTCCCTGCCCTGCCTGTGCCCTGCCCTGCTTC 35969
QY 229 taaccacactctgaaccacagcgctgtggcagggtcccacgtcatgcagactatc 288
Db 35970 CCCCTGCTCCTCTCTTCCCTGCCCTGCCCTCTTTCCTCCCTGCCCTGTCTCCTCCCTGCTC 36029
QY 289 ctctcttgtagcccc 304
Db 36030 CCTCTTCCCTCCCTCCC 36045

RESULT 15
AC006027/c LOCUS
DEFINITION AC006027 200906 bp DNA HTG 22-NOV-1998
*** SEQUENCING IN PROGRESS *** Homo sapiens clone GS114I09; HTGS
phase 1, 2 unordered pieces.
AC006027 ACCESSION
AC006027 NID
g3907480 KEYWORDS
HTG: HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 200906)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 200906)
Waterston,R.H.
Direct Submission
Submitted (22-NOV-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
**
*** WARNING: Phase 1 High Throughput Genome Sequence ***
**
*** This sequence is unfinished. It consists of 2 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
* as runs of N at a convenience only. When sequencing is complete,
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
1 197002: contig of 197002 bp in length
* 197003 197018: gap of unknown length
* 197019 200906: contig of 3888 bp in length.
Location/Qualifiers
1..200906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GS114I09"

BASE COUNT 57076 a 44314 c 42103 g 57397 t 16 others
ORIGIN

Query Match 3.3%; Score 41.2; DB 18; Length 200906;
Best Local Similarity 56.7%; Pred. No. 0.28;
Matches 76; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 106 cttagcttgtcttcctcctcctcctttatttaagaagtccttttatcttcctgc 165
Db 136253 CCTTCCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTC 136194
QY 166 gtacaacactcttcaccccttgcagaccagtccgcltacacctaacccgccgcacactcct 225
Db 136193 TTCTCTGCTTTCCTGCTCCCTGCCCTCTCTCTGCTTCCCTGCCCTGCCCTCTCTCTCCCTCCCTC 136134
226 tgctaccacactct 239
```

Db 136133 CTCCTTCCTTCCT 136120

Search completed: May 14, 1999, 03:36:08
Job time: 5787 sec

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 11:42:35 ; Search time 129.53 Seconds
(without alignments)
1830.042 Million cell updates/sec

Title: US-09-212-270-10

Perfect score: 1260

Sequence: 1 ctccctcaggaggactggaac.....tgcgtgtgtgtagatgagg 1260

Scoring table: IDENTITY_MNC

Searched: 240622 seqs, 94065609 residues

Database: N_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1186.2	94.1	1717	1	V00493	Human tumour necro
2	671.2	53.3	1275	1	V00494	Human tumour necro
3	327.6	26.0	430	1	O59359	Human brain Expres
4	91.6	7.3	317	1	O59350	Human polykystic k
5	41.2	3.3	53577	1	T18551	Human PKD1 locus b
6	41.2	3.3	53577	1	T94108	Human PKD1 gene. H
7	41.2	3.3	53526	1	T94101	Batten disease gen
8	39	3.1	1732	1	T61336	Batten disease gen
9	38.2	3.0	1798	1	T51756	Batten disease gen
10	37.8	3.0	1734	1	T61334	Batten disease gen
11	37.4	3.0	1732	1	T61337	Batten disease gen
12	37.4	3.0	1732	1	T61338	Batten disease gen
13	37.4	3.0	1733	1	T61339	Batten disease gen
14	37.4	3.0	1730	1	T61340	Batten disease gen
15	37.4	3.0	1732	1	T61341	Batten disease gen
16	37.4	3.0	1733	1	T61342	Batten disease gen
17	37.4	3.0	1732	1	T61343	Batten disease gen
18	37.4	3.0	1732	1	T61344	Batten disease gen
19	37.4	3.0	1732	1	T61345	Batten disease gen
20	37.4	3.0	1732	1	T61346	Batten disease gen
21	37.4	3.0	1732	1	T61347	Batten disease gen
22	37.4	3.0	1731	1	T61348	Batten disease gen
23	37.4	3.0	1732	1	T61332	Batten disease gen
24	37.4	3.0	1732	1	T61306	Batten disease gen
25	36	2.9	1978	1	O74448	Distal regulatory
26	35	2.9	9115	1	O74449	myod retroviral ex
27	35.8	2.8	1732	1	T61335	Batten disease gen
28	34.8	2.8	4742	1	O33392	ASM genomic DNA. P
29	34.8	2.8	2347	1	O33390	ASM CDNA, PASM-1FL
30	34.8	2.8	2347	1	O33393	R496L ASM CDNA. Pu
31	34.8	2.8	2344	1	O33384	R496L ASM CDNA. Pu
32	34.8	2.8	2347	1	O33385	L302P ASM CDNA. Pu
33	34.8	2.8	2347	1	T95065	CDNA encoding huma
34	34.8	2.8	2347	1	T95065	CDNA encoding huma
35	34.8	2.8	4742	1	T95070	Human acid sphingo
36	34.8	2.8	2344	1	T95066	CDNA encoding huma
37	34.8	2.8	2347	1	T95067	CDNA encoding huma
38	34.8	2.8	2346	1	T95068	CDNA encoding huma
39	34.6	2.7	1371	1	V18080	Pig liver ribonuc1
40	33.2	2.6	2108	1	V01731	Mouse liver cancer
41	32.8	2.6	1389	1	V21235	Homo sapiens clone
42	32.8	2.6	1717	1	V08170	KMP19 coding seque
43	32.4	2.6	1109	1	V39984	Human T cell surfa

ALIGNMENTS

44 32.4 2.6 1551 1 V58355 Coding sequence fo
c 45 32.4 2.6 641 1 V62167 HSV-2 strain SBS C

RESULT 1
V00493
ID V00493 standard; CDNA; 1717 BP.
AC V00493;
DT 14-APR-1998 (first entry)
DE Human tumour necrosis factor delta encoding CDNA.
KW Human; tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
KW tumour necrosis factor epsilon; T cell proliferation;
KW immune regulation; inflammatory response; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 333..1022
FT /tag= a
FT /product= "Tumour necrosis factor delta"
FT /note= "No stop codon where the given protein ends,
FT but the protein could continue giving a coding
FT region to position 1034 which includes a stop
FT codon"

MO9733902-A1.
PD 18-SEP-1997.
PF 14-MAR-1996; U03774.
PR 14-MAR-1996; WO-U03774.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Gentz RL, Ni J, Yu G;
DR WPI; 97-470810/43.
DR P-PSDB; W37002.
PT Nucleic acid encoding human tumour necrosis factors delta and
PT epsilon - useful for destroying cancer cells, or mediating T cell
PT proliferation, immune regulation and inflammatory responses etc.
PS Claim 10, Fig 1; 81pp. English.
CC The present sequence encodes human tumour necrosis factor (TNF) delta.
CC The TNF can be used to treat patients deficient in this factor
CC (optionally by in vivo expression). It is a ligand for TNF and can
CC destroy (by inducing apoptosis) some transformed cell lines (for tumour
CC treatment); mediate cell activity and proliferation (including T cells
CC to stimulate an immune response to viral, bacterial or parasitic
CC infections), also to eliminate autoreactive T cells in e.g. type I
CC (diabetes), and is functionally linked as a primary mediator of immune
CC regulation and the inflammatory response. A typical application of the
CC protein, its agonists or antagonists is prevention of septic shock,
CC inflammation, cerebral malaria, human immunodeficiency virus activation,
CC graft-host reaction, bone resorption (osteoporosis), rheumatoid
CC arthritis and cachexia, also to promote wound healing (by upregulating
CC cell adhesion) and to regulate haematopoiesis. The protein can also be
CC used to isolate its cognate receptors (or receptor-encoding genes).
CC The CDNA can be used to produce recombinant proteins, and fragments of
CC it to isolate full-length or related sequences, for diagnostic detection
CC of mutations (indicative of disease or susceptibility) and for
CC chromosome identification. Analysis of patient samples for presence of
CC the new protein (e.g. using specific antibodies) can be used
CC diagnostically.
SQ Sequence 1717 BP; 398 A; 447 C; 454 G; 418 T;

Query Match 94.1%; Score 1186.2; DB 1; Length 1717;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1238; Conservative 0; Mismatches 13; Indels 4; Gaps 4;
QY 10 ggagctggaactaattctctgagctggaggggtgaggggtcctcaaggaactg 69
DB 17 GGGACTGGAACCTAATCTCTCTGAGCCTGAGGGGAGGGGTGCAAGACACCTG 76
QY 70 gcccacagagagtgccagagagactaacagatcccttagctgcttctctcctc 129
DB 77 TCCCCAGCAGGAGGAGTGCACAGACCTAACAGTACCTTAGATTGCTTCTCTCCTC 136

130 c-ttttatttcgaagctccttcttattctctctgctgtaaacacctctctctctg 188
 137 ctttttttttttcaagttccttttttttttttttttttttttttttttttttt 196
 189 aacacgtccgtaccccttaaccgcccgcacactctctgtctacccactctgaaacac 248
 197 accactgcccgcacacttaccgcccgcacactctctgtctacacactcttgaanacac 256
 249 agctgtgagaggtt-cccacactatgcccacactctctctctgtctagccccaa 307
 257 agcttttgacaggggttcccacactatgcccacactctctctgtctagccccaa 316
 308 a-ggctccagagcaacatgagggggccagtcagagagccgacactcaagttgcccctg 366
 317 agggcctccagcaacatgagggggccagtcagagagccgacactcaagttgcccctg 376
 367 gttgagttggggggcagctctgaggggccggtgtgtg-catgtgtcagctgacccaa 425
 377 gttgagttggggggcagctctgaggggccggtgtgtgcatgagcctgtgacccaa 436
 426 aacagagctgagagcctcagagagaggtgagccggtgacagggagagggccctc 485
 437 AACAGAGCTGCGAGACCTCAGAGAGAGGTGAGCCGCTGCAGAGAGACAGAGGCCCTC 496
 486 ccaagatgaggagaggtatccctgacagagtcctccgagagcagagttccgagccctga 545
 497 CCAGATGGGGAGAGGTATCCCTGGCAGAGTCTCCGGAGCAGATTCGATGCCCTGGA 556
 546 agcctggagagagagagagagatcccgagaaagagagagagtcacacccaaacagaa 605
 557 AGCCGGGAGAGATGGGGAGAGATCCCGGMAAGAGAGAGAGTGCACGCAAAAACAGAA 616
 606 gaagagagactgtgctcgcagactgttcccatcgaagccacactcaagagatgactcga 665
 617 GAGAGAGACTGTGCTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
 666 tctgacagagtgatgctgacacacagctctctgagcgttggagagagccctacagggccaa 725
 677 TGTGACAGAGGTGATGGGCAACAGCTCTTATGGGCTGGGAGAGGCTACAGGCCCAAG 736
 726 atatgtgtccgaatccagagatgctgaggttatactgtctgtatagcaggtctgttca 785
 737 ATATGTTCCGAAATCCAGAGATGCTGAGATTATCTGCTGATACCGAGTCTCTTTCA 796
 786 agaagctgtccacacagagtgctgctgacagagagagagagagagagagagagagag 845
 797 AGAGCTGACTTTACCACTAGGTGAGGTGCTGCTGAGAGAGGCAAGAGAGAGAGAGAG 856
 846 tctattccgagtgataaagagatgctccacacagagagagagagagagagagagagag 905
 857 TCTATTCGAGTATTAAGAGATGCTCCGACCCGAGCCGAGCCGAGCTACAGAGCTGCTA 916
 906 tagcgcaagtgctctccattacacacagagagagagagagagagagagagagagagag 965
 917 TAGCCAGAGTGTCTTCAATTAACACAGAGGAGATTTGATGATGATCAATATCCCGGAG 976
 966 aagggcgaagacttaacactctccacatgagagagagagagagagagagagagagagag 1025
 977 AAGGCGAAGACTTAACCTCTCTCCACATGAGACCTTCTCGGGGTTTGTGAATCTGTGAT 1036
 1026 gtttataaaagagtgctccagctctggaagagacagagtgagagagagagagagagagag 1085
 1037 GTGTATATAAAGTGGCTCCAGCTTGGAGAGACAGAGGGGTGATACATCTGAGAGAGCC 1096
 1086 aagagctagatataaag 1145
 1097 AAGAGCTGATATTAAG 1156
 1146 gctcccgctctcact 1205
 1157 GCTCCCT 1216
 1206 atctgtctctgtctcccatgagagctccgaattctgtgctgtagatgaggg 1260

1217 ATCTGCTTCTGTTCCCATGAGCTCGAATCTTGTGCTGTGATGAGG 1271

RESULT 2
 V00494
 ID V00494 standard; cDNA; 1275 BP.
 AC V00494;
 DT 14-APR-1998 (first entry)
 DE Human tumour necrosis factor epsilon encoding cDNA.
 KW Human; tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
 KW tumour necrosis factor epsilon; T cell proliferation;
 KW immune regulation; inflammatory response; ss.
 OS Homo sapiens.
 FH key Location/Qualifiers
 FT CDS 2..508
 FT /tag= a
 FT /product= "Tumour necrosis factor epsilon"
 PD MO9733902-A1.
 PD 18-SEP-1997.
 PF 14-MAR-1996; U03774.
 PR 14-MAR-1996; WO-003774.
 PA (HOMA-) HUMAN GENOME SCI INC.
 PI Gentz RL, Ni J, Yu G;
 DR WPI: 97-470810/43.
 DR P-PSDB: W37003.
 PT Nucleic acid encoding human tumour necrosis factors delta and
 PT epsilon - useful for destroying cancer cells, or mediating T cell
 PT proliferation, immune regulation and inflammatory responses etc.
 PS Claim 12; Fig 2; 83pp; English.
 CC The present sequence encodes human tumour necrosis factor (TNF) epsilon.
 CC The TNF can be used to treat patients deficient in this factor
 CC (optionally by in vivo expression). It is a ligand for TNF and can
 CC destroy (by inducing apoptosis) some transformed cell lines (for tumour
 CC treatment); mediate cell activity and proliferation (including T cells
 CC to stimulate an immune response to viral, bacterial or parasitic
 CC infections, also to eliminate autoreactive T cells in e.g. type I
 CC diabetes), and is functionally linked as a primary mediator of immune
 CC regulation and the inflammatory response. A typical application of the
 CC protein, its agonists or antagonists is prevention of septic shock,
 CC inflammation, cerebral malaria, human immunodeficiency virus activation,
 CC graft-host reaction, bone resorption (osteoporosis), rheumatoid
 CC arthritis and cachexia, also to promote wound healing (by upregulating
 CC cell adhesion) and to regulate haematopoiesis. The protein can also be
 CC used to isolate its cognate receptors (or receptor-encoding genes).
 CC The cDNA can be used to produce recombinant proteins, and fragments of
 CC it to isolate full-length or related sequences, for diagnostic detection
 CC of mutations (indicative of disease or susceptibility) and for
 CC chromosome identification. Analysis of patient samples for presence of
 CC the new protein (e.g. using specific antibodies) can be used
 CC diagnostically.
 SO Sequence 1275 BP; 283 A; 343 C; 356 G; 293 T;

Query Match 53.3%; Score 671.2; DB 1; Length 1275;
 Best Local Similarity 93.5%; Pred. No. 5, 9e-194;
 Matches 742; Conservative 0; Mismatches 3; Indels 49; Gaps 2;

468 gggagcag 527
 1 GGGGACAG 60
 528 gagttccgagtgctctgag 587
 61 GAGTTCGATGCTCCGAG 120
 588 gctcaccacaaacag 647
 121 GCTACCCCAAAACAG 141
 648 ctccaagagatgactcagatgtagacagagagagagagagagagagagagagagagagagag 707
 141 -----ATGACTTCGATGAG 192

QY	708	agagctacagggcccaagatagtggtccgaatccagagatgctggagattactctgtta	767
Db	193	AGGCGCTACAGGCCCCAAAGAAATGGTGTCCCAATCCAGGAATGCTGGAGTTATCTCTGA	252
QY	768	tagcagaatccctggtttcaagaacgtgaattccaccatgggtcaaggtgtgtctccggaag	827
Db	253	TAGCCAGGTCCTGTTTAAAGACGTGACTTTCACCAATGGGTCAGAGTGTGTCTCGAAGAAG	312
QY	828	ccaaggaagcgagagagactatctccatgtatataaagaatgtccctccaccagagacg	887
Db	313	CCAAGGAAGGAGGAGACTCTATTTCGATGTATTAAGAAGATGCCCTCCACCGGACCG	372
QY	888	ggctcaacaacagctgtctatagcgaagtgctctccattcaaccaagggagatattcag	947
Db	373	GGCCTACAAACAGCTGCTATAGCCAGGTGCTTCCATTACACCAAGGGAGATATTCTGAG	432
QY	948	tgctaatatcccgaggccaagggcggaacttaacctctctccaatggaactctctgg	1007
Db	433	TGTATATATTCGCCGGCAAGGGGAAACTTAACCTCTCTCCACAAGAAACCTTCTGGG	492
QY	1008	gttcgtcaaacgtgattgattgttataaagtgctgcccaagcttggaaagcagggttgg	1067
Db	493	GTTTGTGAATCTGATTTGTTATTAATAAAGTGGCTCCACGCTTGGAAACCAAGGCTGGG	552
QY	1068	tacatactggaagacagccaagagctgagatataaagaagaggaatgtgcagaaacaga	1127
Db	553	TACATATACGAGACAGCAAGAGCTGATTAATAAAGAGAGGGAATGTGACGAAGAAGA	612
QY	1128	ggatcttctcctgggtttggtcccccgtctcctaactttcccttccattccaccacctta	1187
Db	613	GGCGCTCTCCGGGTTTGGGCTCCCGCTCCCTCACTTTCCCTTTTCATCCCAACCCCTTA	672
QY	1188	gacctt-gatttaccagatctctgtctctgttcccccatggaggtccgaattctgtgt	1248
Db	673	GACTTTGGATTTTACGCAATATTCTGCTTCTGTGTCCCATGAGACTCCGAATTTGGCGTG	732
QY	1247	tggttagatgaggg 1260	
Db	733	TGTGTAGATGAGGG 746	

RESULT 3

Q59349/c

Q59349 standard: cDNA: 430 bp.

Q59349:

16-MAR-1994 (first entry)

Human brain Expressed Sequence Tag EST00378.

Gene transcription product; genetic markers; tagging; in vivo;

transcription; mapping; locations; chromosomes; chromosomal; ss.

Homo sapiens.

MO9316178-A.

19-AUG-1993.

12-FEB-1993: U01294.

12-FEB-1992: U8-837195.

(USSH) US DEPT HEALTH & HUMAN SERVICE.

PA Adams MD, Moreno RF, Venter CJ.

DR WPI: 93-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as

PT markers for human genes transcribed in-vivo, facilitate tagging

PT of most human genes

PS Example 4: Page 160: 500pp: English.

CC The Expressed Sequence Tag was isolated from a human brain cDNA

CC library as part of a large set of ESTs which can be used as markers

CC for human genes transcribed in vivo. They can be used to facilitate

CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosomes, for individual or forensic identification, for mapping

CC locations of disease-associated genes, for identification of tissue

CC type, and for prepn. of antisense sequences, probes and constructs.

CC EST00378 has a "poor" coding probability as evaluated using the

CC coding-region prediction program CRM. See also Q59041-Q61440.

Q0 Sequence 430 BP: 109 A: 113 C: 105 G: 100 T:

Query Match	26.0%	Score 327.6	DB 1	Length 430
Best Local Similarity	97.7%	Pred. No. 6.1e-90		
Matches 341	Conservative	0	Mismatches 7	Indels 1
913	gggtcttcctcatcttacacc-aagsggagatctctgagtcgcatcaattcccgagcaaggc	971		
Db	426	gtttgcttccatttttacccaaaggagatrtttctgactgcatratttccccgcacaaagmc	367	
Qy	972	gaacttaaccctctccacatggaaccttcctggggtttgtgaaactgtgattgttta	1031	
Db	366	gaaacttaactctctccacatggaaccttcctggggtttgtgaaactgtgattgttta	307	
Qy	1032	taaaaagtgtcccccagcttgggaagccaggttggtaatactctggagacagccaaagc	1091	
Db	306	taaaaagtgtcccccagcttgggaagccaggttggtaatactctggagacagccaaagc	247	
Qy	1092	taagatataaaggaaggaatgtgcaagaacagaggaattcttcctgggtttgtctcc	1151	
Db	246	taagatataaaggaaggaatgtgcaagaacagaggaattcttcctgggtttgtctcc	187	
Qy	1152	cgctcccaactttccctttcatcccaacccttagaacttggattttaggatactgt	1211	
Db	186	cgctcccaactttttccctttcatcccaacccttagaacttggattttaggatactgt	127	
Qy	1212	ctctcgttcccatggagctccgaattctctgtgtgtgtatgaagg	1260	
Db	126	cttctgttcccatggagctccgaattctctgtgtgtgtatgaagg	78	
RESULT	4			
ID	059350	standard; cDNA; 317 BP.		
AC	059350:			
DT	16-MAR-1994	(first entry)		
DE	Human brain Expressed Sequence Tag EST00379.			
KW	Gene transcription product; genetic markers; tagging; in vivo; transcription; mapping; locations; chromosomes; chromosomal; ss.			
OS	Homo sapiens.			
PN	W0916178-A.			
PD	19-AUG-1993.			
PF	12-FEB-1993; U01294.			
PR	12-FEB-1992; US-837195.			
PA	(US8) US DEPT HEALTH & HUMAN SERVICE.			
PI	Adams MD, Moreno RF, Venter CJ;			
PT	Enriched oligonucleotides and corresp. sequences - used as			
PT	markers for human genes transcribed in-vivo, facilitate tagging			
PT	of most human genes.			
PS	Example 4; Page 160; 500bp; English.			
CC	The Expressed Sequence Tag was isolated from a human brain cDNA			
CC	library as part of a large set of ESTs which can be used as markers			
CC	for human genes transcribed in vivo. They can be used to facilitate			
CC	tagging of most human genes, for mapping locations of expressed genes			
CC	on chromosomes, for individual or forensic identification, for mapping			
CC	locations of disease-associated genes, for identification of tissue			
CC	type, and for prep. of antisense sequences, probes and constructs.			
CC	EST00379 has a "poor" coding probability as evaluated using the			
CC	coding-region prediction program CRM. See also Q59041-Q61440.			
SD	Sequence 317 BP; 79 A; 71 C; 110 G; 54 T;			
Query Match	7.3%	Score 91.6	DB 1	Length 317
Best Local Similarity	81.5%	Pred. No. 2.1e-18		
Matches 106	Conservative	0	Mismatches 24	Indels 0
Qy	502	tatcctgtcgagagctccgcgagagatctcgaatcccttggaagcctgggaagaatgg	561	
Db	149	TAACCTTAACCTCTTTTCATGACGAGAGATTCCGATGCCCTTGGAAGCTTGGAAGCTGG	208	
Qy	562	gagagatcccgaaaggagagcagtggtgtcaccacaaacagaagaagcagcactgtgc	621	
Db	209	GAGAGATCCCGAAAGGAGAGACAGTGTGTCAACCCAAACAGAAAGAGAGGCTTCCAGG	268	

CC found in one Bd family (L39) of Danish origin.
SQ Sequence 1732 BP; 317 A; 547 C; 465 G; 403 T;

Query Match
Best Local Similarity 3.0%; Score 37.4; DB 1; Length 1732;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 ctagagctcagagagaggtgagccgctgcagggagagagccctccagaatggg 495
DB 630 CAAGGCTTATGATGATGCTAGCGAAGACCAACACAGGCTGTCCTCCACAGATGAG 571
QY 496 gaagggtatccctgcagagaggtctcccgagagcagagttccgagtcctggaagcctggag 555
DB 570 AAAAGCAACCAAGAGCAAGCTTCCAGACAGCAAAATCCACTGACGAGAACCCGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGGG 504

RESULT 12

T61338/c
ID T61338 standard; cDNA; 1732 BP.

AC T61338;
DT 24-OCT-1997 (first entry)
DE Batten disease gene mutation in family L189.
KW Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;
KM gene therapy; ss.
OS Homo sapiens.

PI Brenning MH, Gardiner MR,

PI Mole SE, Taschner PEM;

PI WPI: 97-179265/16.

DR P-PSDB: W13583.

PT Batten disease polypeptide - useful to correct absence of wild type

PT polypeptide, or as agonist to enhance activity of wild type

PT polypeptide.

PS Disclosure: Page -: 94pp; English.

CC This sequence shows a novel mutation of the human CLN3 Batten

CC disease (Bd) gene, corresponding to a C768T nucleotide substn. in

CC CLN3 clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3 gene

CC were identified in 32 Bd patients, 17 of these mutations (T61332-48)

CC being in exon locations. The mutations were identified by PCR

CC amplification of CLN3 exons (see also T61331, T77158-86), single

CC strand conformation polymorphism analysis and direct sequencing.

CC The C768T mutation, in exon 8, converts the Gln codon-221 to a

CC STOP codon (see also W13583). It was found in a Bd family of

CC Italian origin.

CC Sequence 1732 BP; 317 A; 545 C; 465 G; 405 T;

Query Match
Best Local Similarity 3.0%; Score 37.4; DB 1; Length 1732;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 ctagagctcagagagaggtgagccgctgcagggagagagccctccagaatggg 495
DB 630 CAAGGCTTATGATGATGCTAGCGAAGACCAACACAGGCTGTCCTCCACAGATGAG 571
QY 496 gaagggtatccctgcagagaggtctcccgagagcagagttccgagtcctggaagcctggag 555
DB 510 TGTAGGG 504

DB 570 AAAAGCAACCAAGAGCAAGCTTCCAGACAGCAAAATCCACTGACGAGAACCCGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGGG 504

RESULT 13

T61339/c
ID T61339 standard; cDNA; 1733 BP.

AC T61339;
DT 24-OCT-1997 (first entry)

DE Batten disease gene mutation in family L250.

KW Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;

KM gene therapy; ss.

OS Homo sapiens.

PI Brenning MH, Gardiner MR,

PI Mole SE, Taschner PEM;

PI WPI: 97-179265/16.

DR P-PSDB: W13584.

PT Batten disease polypeptide - useful to correct absence of wild type

PT polypeptide, or as agonist to enhance activity of wild type

PT polypeptide.

PS Disclosure: Page -: 94pp; English.

CC This sequence shows a novel mutation of the human CLN3 Batten

CC disease (Bd) gene, corresponding to a 1-bp insertion at base 723 of

CC CLN3 cDNA clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3

CC gene were identified in 32 Bd patients, 17 of these mutations

CC (T61332-48) being in exon locations. The mutations were identified

CC by PCR amplification of CLN3 exons (see also T61331, T77158-86),

CC single strand conformation polymorphism analysis and direct

CC sequencing. The 723insg mutation, in exon 8, results in a

CC frameshift after amino acid G195 in the mutant CLN3 polypeptide

CC (W13584). It was found in a Bd family of UK origin.

CC Sequence 1733 BP; 317 A; 546 C; 466 G; 404 T;

Query Match
Best Local Similarity 3.0%; Score 37.4; DB 1; Length 1733;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 ctagagctcagagagaggtgagccgctgcagggagagagccctccagaatggg 495
DB 630 CAAGGCTTATGATGATGCTAGCGAAGACCAACACAGGCTGTCCTCCACAGATGAG 571
QY 496 gaagggtatccctgcagagaggtctcccgagagcagagttccgagtcctggaagcctggag 555
DB 570 AAAAGCAACCAAGAGCAAGCTTCCAGACAGCAAAATCCACTGACGAGAACCCGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGGG 504

RESULT 14

T61340/c
ID T61340 standard; cDNA; 1730 BP.

AC T61340;
DT 24-OCT-1997 (first entry)

DE Batten disease gene mutation in family L116.

KW Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;
KM gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 138..839
FT mutation /tag- a
FT 695
FT /note- b
PN WO9708308-A1. /note- "695delAG 2 bp deletion in Bd family L116"
PD 06-MAR-1997.
PF 30-AUG-1996; U13896.
PR 31-AUG-1995; US-003030.
PA (GEHO) GEN HOSPITAL CORP.
PA (UYLE-) RIKSUNIV LEIDEN.
PI Breuning MH, Gardiner MR, Gusella JF, Lerner TJ;
PI Mole SE, Taschner PEM;
DR WPI; 97-179265/16.
DR P-PSDB; W13585.
PT Batten disease polypeptide - useful to correct absence of wild type
PT polypeptide, or as agonist to enhance activity of wild type
PT polypeptide.
PS Disclosure; Page -: 94pp; English.
CC This sequence shows a novel mutation of the human CLN3 Batten
CC disease (Bd) gene, corresponding to a deletion of bases 695-696 of
CC CLN3 cDNA clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3
CC gene were identified in 32 Bd patients. 17 of these mutations
CC (T61332-48) being in exon locations. The mutations were identified
CC by PCR amplification of CLN3 exons (see also T61331, T77158-86),
CC single strand conformation polymorphism analysis and direct
CC sequencing. The 695delAG mutation, in exon 8, results in a
CC frameshift after amino acid 5185 in the mutant CLN3 polypeptide
CC (W13585). It was found in a Bd family of Italian.
SQ Sequence 1730 BP; 316 A; 546 C; 464 G; 404 T;

Query Match 3.0%; Score 37.4; DB 1; Length 1730;
Best Local Similarity 55.9%; Pred. No. 0.14;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 cagagcctcagagagaggtgagccgctcagagagagagccctccccaagtgg 495
DB 630 CAAGGCTGTGATGAGATCTTACGAGACCAACACACAGCTGTGCTCCACAGAAATGAG 571
QY 496 gaagggtatccctggcagagctcccgagagagagttccgattccctggagagctggag 555
DB 570 AAAAGGCAACAGAGACGAGCTTCCAGAGCAAAATCCACTGACGAGAACCCGGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGCG 504

RESULT 15
T61341/C
ID T61341 standard; cDNA; 1732 BP.
AC T61341;
DT 24-OCT-1997 (first entry)
DE Batten disease gene mutation in family L285.
KM Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;
KW gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 138..1454
FT mutation /tag- a
FT 1020
FT /tag- b
FT /note- "G1020A missense mutation in Bd family
L285"
PN WO9708308-A1.
PD 06-MAR-1997.
PF 30-AUG-1996; U13896.
PR 31-AUG-1995; US-003030.

PA (GEHO) GEN HOSPITAL CORP.
PA (UYLE-) RIKSUNIV LEIDEN.
PI Breuning MH, Gardiner MR, Gusella JF, Lerner TJ;
PI Mole SE, Taschner PEM;
DR WPI; 97-179265/16.
DR P-PSDB; W13586.
PT Batten disease polypeptide - useful to correct absence of wild type
PT polypeptide, or as agonist to enhance activity of wild type
PT polypeptide.
PS Disclosure; Page -: 94pp; English.
CC This sequence shows a novel mutation of the human CLN3 Batten
CC disease (Bd) gene, corresponding to a G1020A nucleotide change in
CC CLN3 cDNA clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3
CC gene were identified in 32 Bd patients. 17 of these mutations
CC (T61332-48) being in exon locations. The mutations were identified
CC by PCR amplification of CLN3 exons (see also T61331, T77158-86),
CC single strand conformation polymorphism analysis and direct
CC sequencing. The G1020A mutation, in exon 11, results in an E295K
CC amino acid change in the mutant CLN3 polypeptide (W13586). It was
CC found in a Bd family (L285) of Finnish origin.
SQ Sequence 1732 BP; 318 A; 546 C; 464 G; 404 T;

Query Match 3.0%; Score 37.4; DB 1; Length 1732;
Best Local Similarity 55.9%; Pred. No. 0.14;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 cagagcctcagagagaggtgagccgctcagagagagagccctccccaagtgg 495
DB 630 CAAGGCTGTGATGAGATCTTACGAGACCAACACACAGCTGTGCTCCACAGAAATGAG 571
QY 496 gaagggtatccctggcagagctcccgagagagagttccgattccctggagagctggag 555
DB 570 AAAAGGCAACAGAGACGAGCTTCCAGAGCAAAATCCACTGACGAGAACCCGGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGCG 504

Search completed: May 14, 1999, 11:42:53
Job time: 5282 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 19:44:17 ; Search time 1016.5 Seconds
(without alignments)
1890.836 Million cell updates/sec

Title: US-09-212-270-10

Perfect score: 1260

Sequence: 1 cctctcaggagcgcgcgaac.....tcgcgtgtgtgtatgaggg 1260

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database :

EST.*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: gb_est1:*
11: gb_est2:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: gb_est6:*
16: gb_est7:*
17: gb_est8:*
18: gb_est9:*
19: gb_est10:*
20: gb_est11:*
21: gb_est12:*
22: gb_est13:*
23: gb_est14:*
24: gb_est15:*
25: gb_est16:*
26: gb_est17:*
27: gb_est18:*
28: gb_est19:*
29: gb_est20:*
30: gb_est21:*
31: gb_est22:*
32: gb_est23:*
33: gb_est24:*
34: gb_est25:*
35: gb_est26:*
36: gb_est27:*
37: gb_est28:*
38: gb_est29:*
39: gb_est30:*
40: gb_est31:*
41: gb_est32:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	481.4	38.2	495	24	AA405973	AA405973 zu57g07.r
2	423	33.6	441	21	AA477087	AA477087 zu35f06.r

RESULT	LOCUS	DEFINITION	ACCESSION	NID	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	AA405973	zu57g07.r1 Soares ovary tumor NBH07 Homo sapiens	495 bp mRNA	EST	09-NOV-1997	742140	human.	Primates; Catarrhini; Homindae; Homo.	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810	unpublished (1997)	

ALIGNMENTS

3	422.2	33.5	438	21	AA291913
C	421.4	33.4	479	21 <td>AA481449</td>	AA481449
C	409.8	32.5	439	20 <td>AA443577</td>	AA443577
C	409.4	32.5	411	26 <td>AA779357</td>	AA779357
C	394.4	31.3	462	31 <td>AA149460</td>	AA149460
C	387.4	30.7	467	19 <td>AA056924</td>	AA056924
C	379	30.1	435	21 <td>AA293679</td>	AA293679
C	372	29.5	428	28 <td>AA394070</td>	AA394070
C	369.4	29.3	398	21 <td>AA292358</td>	AA292358
C	365	29.0	398	21 <td>AA292358</td>	AA292358
C	360.2	28.6	583	26 <td>AA790394</td>	AA790394
C	357	28.3	473	31 <td>AA271780</td>	AA271780
C	349.4	27.7	448	23 <td>AA128183</td>	AA128183
C	345	27.4	401	24 <td>AA405974</td>	AA405974
C	343.2	27.2	356	24 <td>AA594343</td>	AA594343
C	341	27.1	470	23 <td>AA298067</td>	AA298067
C	336.6	26.7	447	26 <td>AA782470</td>	AA782470
C	332.4	26.4	390	27 <td>AA600311</td>	AA600311
C	327.6	26.0	430	11 <td>W78230</td>	W78230
C	322	25.6	378	21 <td>AA477507</td>	AA477507
C	304.4	24.2	377	23 <td>AA1216415</td>	AA1216415
C	303	24.0	303	18 <td>AA361896</td>	AA361896
C	300	23.8	405	19 <td>AA057069</td>	AA057069
C	295	23.4	320	18 <td>AA337828</td>	AA337828
C	292	23.2	330	19 <td>AA401109</td>	AA401109
C	289.4	23.0	347	31 <td>AA1250901</td>	AA1250901
C	283.2	22.5	465	24 <td>AA476861</td>	AA476861
C	278.2	22.1	368	18 <td>AA360555</td>	AA360555
C	261.6	20.8	294	22	C75095
C	182.4	14.5	240	31 <td>AA1244559</td>	AA1244559
C	175.4	13.9	239	13 <td>N91002</td>	N91002
C	169.2	13.4	232	13 <td>W20131</td>	W20131
C	148	11.7	224	24 <td>AA405918</td>	AA405918
C	140.2	11.1	347	19 <td>AA401110</td>	AA401110
C	137	10.9	225	19 <td>AA293091</td>	AA293091
C	136.4	10.8	194	22 <td>AA526362</td>	AA526362
C	135.4	10.7	193	22 <td>AA526281</td>	AA526281
C	115	9.1	171	19 <td>AA293481</td>	AA293481
C	115	9.1	171	24 <td>AA405795</td>	AA405795


```

|||||
Db 259 GGAACCTTCTCGGCGTTTGAACAGCTATGTGTATTAATAAAGCGCTCCAGCTTGG 318
Oy 1054 aagaccagggtggtacatactctgagacagcagaagctgagatataaaggagagaa 1113
Db 319 AAGACCAGGAGGTGGTACATCTAGAGACAGCCAGAGCTGAGATATTAAGAGAGAGAA 378
Oy 1114 tctgagagagagagagagatctctcctggtgttggctcccgctcctcctccttcttc 1173
Db 379 TGTGAGGAACAGAGAGCATCTTCTCGGTTTGGCTCCCGTTCCTCATCTTCCCTTTTC 438
Oy 1174 att 1176
Db 439 ATT 441

RESULT 3
AA291913 438 bp mRNA EST 08-AUG-1997
LOCUS zt38h12.r1 Soares ovary tumor NBHOT Homo sapiens CDNA clone 724679
DEFINITION 5', mRNA sequence.
ACCESSION AA291913
NID 91939891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 438)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, B., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
WashU-Merck EST Project
unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1277 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 426.
Location/Qualifiers
1..438
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGGTTTATTTTATTTT 3'],"
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
<1..>438

MRNA
BASE COUNT 78 a 147 c 106 g 107 t
ORIGIN
Query Match 33.5%; Score 422.2; DB 21; Length 438;
Best Local Similarity 99.1%; Pred. No. 8.3e-110;
```

```

Matches 435: Conservative 0; Mismatches 3; Indels 1; Gaps 1:
Oy 1 cttctagagaggagctggaacctaatctctctgagctgagggagggtgaggtctcaag 60
Db 1 CTTCCTAGAGAGGAGCTGGAACCTAATCTCTCTGAGGCTGAGGAGGTGAGGCTCAAG 60
Oy 61 gaaagctgccccacagagagagtgccagagagcactaaagtaaccttgccttcc 120
Db 61 GCAAGCTGGCCCCACAGAGAGGTGCCAGAGACACTAACAGTACCTTAGCTTGCTTTC 120
Oy 121 tctctcctctcttatttctcaagttccttcttattctcctctgctgaacacctcttc 180
Db 121 TCTCTCCTCTCTTATTATTTCAGTTCCTTTTATTTCCTCTGCTAACACTTCTTC 180
Oy 181 cctctgcaacactgcccgtacaccttaaccgcccgcacactctgtctacccactct 240
Db 181 CCTTGTGACACAGTGGCCGCTTACCTTACCAGCCGCCACACTCTGTGTAACCACTCTT 240
Oy 241 gaaacacagctgttggcagaggtccccaagctatgccaagctcactctctctgtctag 300
Db 241 GAACACACAGCTGTGGCGAGGCTCCCACTCATGCTCATCTCTTCTTCTGTAG 300
Oy 301 cccccaagagcctcgaaggacatgaggggcccgatcagagagccggacactcagttgc 360
Db 301 CCCCAGAGGCTCCAGG-AACATGGGGGGCCAGTCAAGAGCCGGCACTCTCACTTTC 359
Oy 361 cctctggttgaattgaggagcagctctgaggggcgctgttgcattgagctcagctgacc 420
Db 360 CCTCTGTTGAGTTGGGGGCGAGCTCTGGGGGCGGTGCTTGATGCTGTGCTGAC 419
Oy 421 caacaacagagctgcaga 439
Db 420 CAACAACAGAGCTGCAGA 438

RESULT 4
AA481449 479 bp mRNA EST 08-AUG-1997
LOCUS zva5col.s1 Soares ovary tumor NBHOT Homo sapiens CDNA clone 756576
DEFINITION 3', mRNA sequence.
ACCESSION AA481449
NID 92211001
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 479)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kucaba, T., Lacey, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B.,
White, Y., Wylie, T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
unpublished (1997)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 468.
Location/Qualifiers
1..479
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGGTTTATTTTATTTT 3'],"
```

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.

/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>479)
BASE COUNT 122 a 128 c 121 g 108 t
ORIGIN

Query Match 33.4%; Score 421.4; DB 21; Length 479;
Best Local Similarity 99.8%; Pred. No. 1.4e-109;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 838 caggagactctatccgagtataagaagatcccccacggagcgagcctacaac 897
DB 479 CAGAGACCTATTCCCATATAGAGATAGCCCTCCACCGGACCGGCTTACAC 420
QY 898 agctgctatagcgaagtgctctccattacacaaaggagatctcgaagtcataat 957
DB 419 AGCTGATAGCGCAGGTGCTTCATTACACCAAGGAGATATTCGAGTGCATATT 360
QY 958 ccccgagcaaggcggaacttaacctctccacatggaacctctcggtgttgataa 1017
DB 359 CCCCCGCAAGGGCGAAGCTTAACCTCTCCACATMAACCTTCCTCGGGGTTGTGAAA 300
QY 1018 ctgtgattgtgtataaaggctcccgcttggaagacagagtggttatactagg 1077
DB 299 CTGTGATGTGTTATATAAAGTGGCTCCAGCTTGGAAGACAGGTTGACTACTGG 240
QY 1078 agacagcccaagagctgagatataaaggagaggaatgltcaggaagaaggagcattcc 1137
DB 239 AGACAGCCAGAGCTGAGTATATAAGAGAGAGGAATGTCCAGGAACAGAGGACTCTCC 180
QY 1138 tgggtttggtcccgctctcaactttccctttcaatccacacccctgaacttgatt 1197
DB 179 TGGTGTGGCTCCCGCTCCCTCACTTTCCCTTTTCATTCACACCCCTGAGACTTTGATT 120
QY 1198 ttccgagatctgctctgttcccatgagctccgaattcttgctggttataagta 1257
DB 119 TTTAGGATATCTTCTTCTTCTTCCCATGAGCTCCGAATTTTCGTGTGTAGATGA 60
QY 1258 ggg 999 1260
DB 59 GGG 57

RESULT 5
AA443577/c 439 bp mRNA EST 03-JUN-1997
LOCUS 2w35801.s1 Soares ovary tumor NBHOT Homo sapiens cDNA clone 771240
DEFINITION 3', mRNA sequence.
ACCESSION AA443577
NID 92156252
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, I., Wylie, T., Waterston, R., and Wilson, R.
TITLE Washu-Merck EST Project 1997
JOURNAL Unpublished (1997)

COMMENT

Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -41m3 fwd. Et from Amersham
High quality sequence stop: 401.
Location/Qualifiers

FEATURES

source 1..439
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAGTGGAGCGCGCGGTTTTTTTTTTT 3']"
double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>439)
/db_xref="GDB:598116"
BASE COUNT 87 a 127 c 116 g 109 t
ORIGIN

Query Match 32.5%; Score 409.8; DB 20; Length 439;
Best Local Similarity 99.3%; Pred. No. 2.8e-106;
Matches 422; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 513 gaagctccggagcagagctcgcgcttggaagcctggaagatggaagatcccg 572
DB 439 GAGTCTCCGGAGCAGAGTTCGATCCCTCGAAGCCTGGGAAGTGGGAGATCCCG 380
QY 573 gaaagaagagcagctgtctaccccaaaacagaagaagcagcctgtctgcactggt 632
DB 379 GAAAGAGAGCAGTGTCTCAACCAAAACAGAGAGCAGCACTGTCTGCACCTGTT 320
QY 633 ttccattacgcccactcccaagatgactccgagtgcagagagtgatgtgcaaccagc 692
DB 319 TCCCATTTAAGCCACTCCAGAGATACCTCGATGTGACAGAGATGTGGCAACACAG 260
QY 693 tcttaagcgttggaagagcctacagcccaagatatagtgtgtccgaatccagatctg 752
DB 259 TCTTAGCGCGGAGAGG-CCCTACAGGCCCAAGAGATGTGTGCCAATCCAGATCTGG 201
QY 753 agttatctgctgtatagcgaagctctgtttcaagaagtgactttcacatggtcagt 812
DB 200 AGTTATCTGCTGTATAGCAGGTCTGTTCAGAGCGTACTTTCCATCATGGGTAGGT 141
QY 813 ggtgtctcagaagagcgaagagcagagagactctatccgagtataagaagatgac 872
DB 140 GGTGTCTCGAAGAGCCAGAGAGGAGAGAGTATTTCCGATGTATAGAGATGTGCG 81
QY 873 ctccacccgagacgggcttacacagctgctatagcaggtgtcttcattacacca 932
DB 80 CTCACACCGGAGCGGCGCTACACAGCTGCTATACGAGAGTCTTCATTATACCT 21
QY 933 aaggag 937
DB 20 CCGGG 16

```

RESULT 6
AA779357/c
LOCUS AA779357 411 bp mRNA EST 05-FEB-1998
DEFINITION Z163d06.s1 Soares fetal liver spleen INFLS S1 Homo sapiens CDNA
clone 454955 3', mRNA sequence.
ACCESSION AA779357
NID 92838688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 411)
AUTHORS Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritzman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Scheinberg,K., Stepec,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
Mashu-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: eschwatson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 399.
Location/Qualifiers
1. .411
/organism="Homo sapiens"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AAGTGAAGATATATTAAGATCTTTTATTTTATTTT 3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bernaldo."
/db_xref="GDB:1391311"
/db_xref="taxon:9606"
/clone="454955"
/clone_1lb="Soares fetal liver spleen INFLS S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 96 a 116 c 92 g 107 t
ORIGIN
Query Match 32.5%; Score 409.4; DB 26; Length 411;
Best Local Similarity 99.8%; Pred. No. 3.6e-106;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 711 cctaaagcccaagatgtgtgtcgaatccagatgtcgtgaattatctctatag 770
Db 411 CCTAAGGCCCAAGATATGTGTCTCGAATCCAGATGCTGAGTTTATCTGCTATAG 352
Oy 771 ccaagtcctgttccaagcgtacttcaaccatgggtcaggtgtctcgaagagcca 830
Db 351 CCAGTCTCTTTCAAGACGTACTTTCACCATGGGTCAAGGTGTCTCCGAAGGCCA 292
Oy 831 aggaagcagagactctattccagtataaagaatgacctcccccacccgagccggc 890
Db 291 AGGAAGCAGGAGACTCTATTCGATGATAGAGATAGCCCTCAACCCGAGCCGGC 232
Oy 891 ctacacagcgtctctagcgaggtgtcttcattacccaagggatattctaggt 950

```

```

Db 231 CTACACAGCTGCTTACCGCAGTGTCTTCATTATACCAAGGAGATTTCAGTCT 172
Oy 951 caataccccggggaagggcgaacttaacctctccacatggaaaccttcctgggt 1010
Db 171 CATAATTCGCCGGGCAAGCGCAACTTAACTCTCTCCACATGGAACCTTCTCGGGGT 112
Oy 1011 tctgaactctgtatgtcttataaaagtgtctcccaacttggaagaccaggtgtgac 1070
Db 111 TGTGAACGTGTGATGTGTATTAAGTGTCTCCACTTGGAGACAGGTGTGATAC 52
Oy 1071 ataccgagacagcccaagctgagtatataaagagagaggaatgtcagc 1121
Db 51 ATACTGGAGACAGCCAGAGAGCTGAGTATATTAAGAGAGGAGATGTGAGG 1
RESULT 7
A1149460/c
LOCUS A1149460 462 bp mRNA EST 28-OCT-1998
DEFINITION qc78b07.x1 Soares Placenta_8to9weeks_2NBP8to9W Homo sapiens CDNA
clone IMAGE:1715701 3', mRNA sequence.
ACCESSION A1149460
NID 93677929
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 462)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 607 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 457.
Location/Qualifiers
1. .462
/organism="Homo sapiens"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCGCATTTTATTTTATTTT 3'] ,
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bernaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1715701"
/clone_1lb="Soares Placenta_8to9weeks_2NBP8to9W"
/dev_stage="two placenta: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 114 a 126 c 120 g 102 t
ORIGIN
Query Match 31.3%; Score 394.4; DB 31; Length 462;
Best Local Similarity 99.7%; Pred. No. 6.7e-102;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 865 agatgcccctccacccgagccgacctacaacaactgtctatagcgaggtcttccat 924
Db 462 AGTAGCCCTCCACCGGAGCGGCGCTACAAACACTGCTATAGCGAGGTCTTCCAT 403
Oy 925 ttacacaaagggatattctaggtctcaattccccgggcaagggcgaactaacctc 984

```


Seq primer: -41m13 fwd. ET from Amerisham
High quality sequence stop: 398.
Location/Qualifiers

FEATURES

source
1..435
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TCTTACCACTCTGAAGTGGAGCGCGCGGCTTTTCTTTTCTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaudo."
/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBH09"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>435)
/db_xref="GDB:5936317"
BASE COUNT 112 a 121 c 107 g 95 t
ORIGIN

Query Match 30.1%; Score 379; DB 21; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.6e-97;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 882 gggccggggcctacaagcgcgtatagcgaagtgcttcattcaaccaaggagat 941
DB 435 GGACCGGGCCCTACAAAGCTGCTATAGCGAGGTCCTTCATTATACCAAGGGATAT 376
QY 942 tctgaagtacataatcccccggcgaagggcgaactaacctctccacatgaaacct 1001
DB 375 TCTGAGTGTCTATATTCCTCCGGGCAAGGGCAACTTAACCTCTCTCCACATGGAACTT 316
QY 1002 ccttggggtttgtgaaactgtatgtgtataaaagtggtcccaagcttgaagaccag 1061
DB 315 CCGGGGGTTGTGAACACTGATGTGTATATAAAGTGGCTCCACACTTGGAAAGACAG 256
QY 1062 ggtgggtacactatgagagaagcccaagagctgagatataaagagagggaaatgtagcagg 1121
DB 255 GGTGGGTACTACTGAGAGCAAGCAAGCTGATATATTAAGAGAGGGAATGTGCAGG 196
QY 1122 aaagaggaactcttcctggtttgggtcccggtccctcaatttcccttcaattccac 1181
DB 195 AACAGAGGCACTCTTCTGGTGGTGGCTCCCGCTCCACATTTTCCCTTTTCATTTCCAC 136
QY 1182 cccctagacttgatttaagatatctgtctgttcccaatgagagctccgaattctt 1241
DB 135 CCCCTAGACTTTGATTTTACGATATCTTCTCTGTCCCATGAGAGTCCGAAATTCCT 76
QY 1242 ggcgtgtgtgtagatgagg 1260
DB 75 GCCTGTGTGTAGATGAGG 57
RESULT 10
AA394070/c 428 bp mRNA EST 12-AUG-1997
LOCUS 2t53f04.s1 Soares ovary tumor NBH09 Homo sapiens cDNA clone 725983
DEFINITION 3', mRNA sequence.
ACCESSION AA394070
NID 92047040
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 428)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Matra, M., Mattin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE JOURNAL
COMMENT Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 454 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amerisham
High quality sequence stop: 393.
Location/Qualifiers

FEATURES

source
1..428
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TCTTACCACTCTGAAGTGGAGCGCGCGGCTTTTCTTTTCTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaudo."
/db_xref="taxon:9606"
/clone_lib="725983"
/clone_lib="Soares ovary tumor NBH09"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>428)
/db_xref="GDB:5937985"
BASE COUNT 112 a 117 c 105 g 94 t
ORIGIN

Query Match 29.5%; Score 372; DB 21; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.5e-95;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 889 gctacaagcgcgtatagcgaagtgcttcattcaaccaaggagatattagat 948
DB 428 GCCTACAAAGCTGCTATACGCAAGGTCCTTCATTATACCAAGGGATATTCGAT 369
QY 949 gtacataatcccccggcgaagggcgaactaacctctctccacatgaaaccttctggg 1008
DB 368 GTATATATTCCTCCGGGCAAGGGCAAACTTAACCTCTCTCCACATGAACTTCCTGGG 309
QY 1009 tttgtgaaactgtatgtgttataaaagtggtcccaagcttgaagaccagggtggt 1068
DB 308 TTTGTGAACCTGTGATGTGTATATAAAGTGGCTCCACACTTGAAGACCAAGGTGGGT 249
QY 1069 acatactggagacagccaagagctgatatataaagagggagatgtgcaagaaagag 1128
DB 248 ACATACTGGAGACGCCAAGAGCTGATATATAAGAGAGGAGATGTGACGAAACGAG 189
QY 1129 gcatcttcctgggttgggtcccggtccctcaatttcccttcaattccacaccttag 1188
DB 188 GCATCTTCCTGGGTTGGCTCCCGCTCCACATTTTCCCTTTTCATTTCCACCCCTAG 129
QY 1189 acttgatttacgabatctgtcttctgttcccaatgagagctccgaattcttgcgtgtg 1248
DB 128 ACTTGTATTTACGATATCTTCTCTGTCCCATGAGAGTCCGAATCTTGCGGTG 69
QY 1249 tttagatgagg 1260
DB 68 TGTAGATGAGG 57

RESULT 11
AA916330/c
LOCUS
DEFINITION
AA916330
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA916330 395 bp mRNA EST 19-MAY-1998
O180B09.s1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1473305 3',
mRNA sequence.
AA916330
g3055722
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 395)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:
www-bio.liml.gov/bdrip/image/image.html

Insert Length: 1318 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 385.
Location/Qualifiers
1. 395
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_image="1473305"
/clone_lib="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/lab_host="DH10B"

BASE COUNT 107 a 106 c 95 g 87 t
ORIGIN

Query Match 29.3%; Score 369.4; DB 28; Length 395;
Best Local Similarity 99.7%; Pred. No. 8.2e-95;
Matches 370; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

890 cctcaacagctgtctatagcgaggtgtcttcatttccacaaaggagatattcgtatg 949
|||||
Db 395 CCTACACAGCTGCTATAGCGCAGGTGCTTCATTACACCAAGGATATTCTGAGTG 336
950 tcatatccccggcgcaaggaggaaacttaactctctccacatggaactctcggagt 1009
|||||
Db 335 TCATATATTCCTCCGGGCAAGGGGGAACCTTAACCTCTCTCCACATGGAACCTTCCCTGGGGT 276
1010 ttgttaacctgtatgtgttataaaagtgtgtcccgacttggagaagccaaagtggtgta 1069
|||||
Db 275 TTGTAAACTGTGATTTGTATTAAAAAGTGGCTCCGACCTTGGAAAGACCAAGGTGGGTA 216
1070 catactgagacagccaaagagctgagtatataaaggagaggaagtaagtcgaagaaacagag 1129
|||||

Db 215 CATACTGAGACAGCCAGACGCTAGTATATAAGAGAGAGGAATGTGCAGAAACAGAGG 156
Qy 1130 catcttcgtgtgtgtgtcccgcttcctacttcccttccattccacccccaga 1189
|||||
Db 155 CGCTCTCCGGGTTGGTCCCGCTTCCTACCTTTCCCTTTTATTCCACCCCTGAGA 96
Qy 1190 ctctgatttaccgatatctgtctctgttcccatgagctccgaattcttgcgtgt 1249
|||||
Db 95 CTTGATTTTACGATATNCTTCTGTTGTTCCCATGAGAGCTCGAATCTCTGCGTGTGT 36
Qy 1250 gtatagtagagg 1260
|||||
Db 35 GTAGATGAGGG 25

RESULT 12
AA292358
LOCUS
DEFINITION
AA292358
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA292358 398 bp mRNA EST 08-AUG-1997
zt51f04.r1 Soares ovary tumor NDHOT Homo sapiens cDNA clone 725887
5', mRNA sequence.
AA292358
g1940338
EST.
human.
Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;
Homo.
1 (bases 1 to 398)
Haller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E.,
Waters, R., Williamson, A., Wohlmann, P. and Wilson, R.
Washu-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (infoimage.liml.gov) for further information.
Insert Length: 1205 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 370.
Location/Qualifiers
1. 398
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCGGTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo."
/db_xref="taxon:9606"
/db_xref="725887"
/clone_lib="Soares ovary tumor NDHOT"
/sex="female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/db_xref="GDB:5937862"
/db_xref="GDB:5937862"

BASE COUNT 78 a 129 c 113 g 72 t 6 others
ORIGIN

Query Match 29.0%; Score 365; DB 21; Length 398;
Best Local Similarity 95.7%; Pred. No. 1.5e-93;

Matches 382; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 169 acaacttttccctctgacacacgcccgtaccccttaaccgccccgacactcttcg 228
|||||
Db 1 ACAACCTTCTCCCTCTCCTCACCACCTGCCGTACCTTACCCGCCGCCACCTCTTGC 60
QY 229 taaccacacttgaaccacacagctgttgcaagggtcccccagctcagccacatctc 288
|||||
Db 61 TACCCACACTTGAACACACAGCTGTGGCAGGGGTCCCAAGCTCATGCCACTCATTC 120
QY 289 cttctctgtagcccccacaaaggccctccagcaacatggtggggccagtcagagccggc 348
|||||
Db 121 CTTTCTTGTAGCCCCCAAGCGCTCCAGGACATGGGGGCCCAAGTCAGAGAGCCGGC 180
QY 349 actctcagttgcctctgtgttggtgtgggggagctctcgtggggccgtgtgttcacg 408
|||||
Db 181 ACTCTCAGTTGCCCTCTGTGTTGAGTTGGGGGGCAGCTCGGGGGCCGTGGTGTGCATG 240
QY 409 gtccagctgacccaacaacagagctgcagagcctcagagagagagtgagagctgcag 468
|||||
Db 241 CTGACACTGACCCACACACAGAGCTGCAGACCTCAGGAGAGAGGTGACCCGGCTGCAG 300
QY 469 gggacagagagagccctccagaaatgggagaggtatccctgagagagctcccgagagcag 528
|||||
Db 301 GGGACAGAGAG--CCTNCCAGATGANNAGGATATCCCTGNCAGAGTCTCCCGAGAGCAG 359
QY 529 agtccgagtcctcgtgagagcctgagagagagagagagagagagagagagagagag 567
|||||
Db 360 AGTCCGATGCCCTGTGAGCTGGAGAGTGGGGAGAGA 398

RESULT 13

AA790394 583 bp mRNA EST 06-FEB-1998
LOCUS vw04a1.r1 Soares mouse mammary gland NBMKG Mus musculus cDNA clone
DEFINITION 1230908 5', mRNA sequence.
ACCESSION AA790394
NID g2850514
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Rodentia: Sciuromorphi; Muridae: Murinae; Mus.
1 (bases 1 to 583)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:656500
Putative full length read
vector to vector length is 891
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 499.

FEATURES
source 1..583
location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker. Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGGCGCGCGAGATGGTTTATTTTATTTTATTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonafide. "
/db_xref="taxon:10090"
/clone_lib="Soares mouse mammary gland NBMKG"
/clone_1ib="1230908"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 140 a 133 c 187 g 122 t 1 others
ORIGIN

Query Match 28 6%; Score 360.2; DB 26; Length 583;
Best Local Similarity 79.5%; Pred. No. 3.6e-92;
Matches 464; Conservative 0; Mismatches 114; Indels 6; Gaps 3;

QY 308 aggcctccagagacacatggggggcccgacagagagcggcaactcagttgccccttg 367
|||||
Db 5 AAGCTCCAGGCCACATGGGGGGCTCAGTCAGAGGCCAGCCCTTTCGTTGCTTTGG 64
QY 368 ttgagttggggggcagctctggtggggcggtgtgtgc--atggttcagctgacacacaa 426
|||||
Db 65 TTGAGTTGGGGGCGCATTTCTGGGGGCTGTGACTGTGTGTGCTGCACTACTGATCCAAACG 124
QY 427 acagagctcagagagcctcagagagaggtgagccggtcagagggagacagagagagag 486
|||||
Db 125 ACAGAGCTCTCAACCTCAAGCGGAGGAGGTGAGCGGCTCAGCGGAGTGGAGGCTTCC 184
QY 487 cagaaatggggaagggtatccctcgtgagagagctcccgagcaagagttcagagagagag 546
|||||
Db 185 CAGAGCAGGAGGAGAGGCCCATG--CAGAGCTTGGAGAGAGAGTCTGTGTGCTTGGAA 242
QY 547 ggcctggagaaatgggagagatcccgaaagagagagagagagagagagagagagagag 606
|||||
Db 243 GCCTGGAAGAGTGGGGCGAATCTCGAGAGAGAGAGAGTACTACCCAGAGCACAAG 302
QY 607 aagcagcactctgctcagctcagctgtgtccatcaagcagcctcagagagagagagag 666
|||||
Db 303 AAGAACCACTCACTCTCTGATCTTCCAGTTAACATTACTCTCAAG--GACTCTGAC 359
QY 667 gtagcagaggtgtagtgggaacacagctcttaggcgtgggagagagagagagagagag 726
|||||
Db 360 GTACAGAGGTGATGTGGCAACCACTTAGGCGGTGGGAGAGGCTTGGAGGCCAGGGA 419
QY 727 tatggtccgaatccagagatgctgaggttatctgtatagcagagctcgtttcaa 786
|||||
Db 420 GACATGTGACGAGTCTGGGACACTGGAATTTATCTCTATATGTAGAGCTCTTTCA 479
QY 787 gacgtgacttcaacacatgggtcaggtgtgtctcgagaaagggcagagagagagagag 846
|||||
Db 480 GATGTGACTTTTCAATGAGGTGAGTGTATCTCGGGAAGAGACAGGAGAAAGAGAACT 539
QY 847 ctattccgattgataagagatgacccctccacccgagagagagagagagagagagag 890
|||||
Db 540 CTATTCGATGATACAGAGTATGCTTGTGATCTGATCTGACCGTCG 583

RESULT 14

AI271780/c 473 bp mRNA EST 17-NOV-1998
LOCUS AI271780
DEFINITION qj82f03.x1 NCI_CGAP_K1d3 Homo sapiens cDNA clone IMAGE:1865981 3',
ACCESSION AI271780
NID q3890947
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 440.
Location/Qualifiers
1. 473
/organism="Homo sapiens"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."
/db_xref="taxon:9606"
/db_xref="IMAGE:1865981"
/clone_1lb="NCI-CGAP_K1d3"
/lab_host="DH10B"
BASE COUNT 120 a 130 c 119 g 103 t 1 others
ORIGIN

Query Match 28.3%; Score 357; DB 31; Length 473;
Best Local Similarity 97.4%; Pred. No. 2.8e-91;
Matches 405; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

QY 845 cctcattccgagtataagaatgctccctccacccgagccggtctacacagctgct 904
DB 467 cctcattccgagtataagaatgctccctccacccgagccggtctacacagctgct 904
QY 905 atagcgcagcgtgcttcattacacacagcggatattctgagtataattcccgagg 964
DB 410 ATACGCGAGGtGtGTCCTTTA--ACCAAGGGGATATTCTAGTGTCTAATTATTCGCCGG 353
QY 965 caagcggcgaacttaacctctccacatggaacctctcgtgggtttgtgaactgtgat 1024
DB 352 CAAGGGCGAAACTTAACCTCTCCACATGAGACCTTCTGGGTTTGTGAATGTGAT 293
QY 1025 tctgttataaagttgctcccaagcttggaaagcagggtgggtgatactggaacagc 1084
DB 292 TGTGTATTAATAAAGTGGCTCCACGCTTGGAAGACGAGGTGGTACATACGAGACAGC 233
QY 1085 caaagcgtgagtataaagaaggaggatgtgcaagcagcaggcatcttcctgggttt 1144
DB 232 CAAGAAGCTGAGTATATAAAGAGAGGAGATGTGAGAGAACAGAGGCGCTTCCTGGGTTT 173
QY 1145 ggtcccccgttccctcattcccttccattcccaaccctagactgttttaagcga 1204
DB 172 GGGCCCGCGTTCACACTTTCCTTTTCATTTCCACCCCTAGACTTTGATTTTACGGA 113
QY 1205 tacccttcctgttcccaatgagagctcgaattcttgcgtgtgtatgataagag 1260
|||||

DB 112 TACTTCTCTGTGTCCCATGAGAGCTCCGATTCCTGCTGTGTAGATGAGAG 57
RESULT 15
AI281823/c 448 bp mRNA EST 23-NOV-1998
LOCUS q587a03.x1 NCI-CGAP_Col4 Homo sapiens cDNA clone IMAGE:1962220 3',
DEFINITION mRNA sequence.
ACCESSION AI281823
NID 93920056
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 448)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1. 448
/organism="Homo sapiens"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 kb. Life Technologies catalog #:
11531-019"
/db_xref="taxon:9606"
/clone_1lb="NCI-CGAP_Col4"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 113 a 123 c 114 g 97 t 1 others
ORIGIN

Query Match 27.7%; Score 349.4; DB 23; Length 448;
Best Local Similarity 99.2%; Pred. No. 4e-89;
Matches 361; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 897 cagctgtcagaagcagcgttcttcattacacacagcggagatattcgtgtccta 956
DB 448 CAGCTGCTATAGGCGCAGGtGTCCTTTCATTACACCAAGGGGATATTCGTGATCATAT 389
QY 957 tccccggcgaaggcgaacttaacctctccacatggaacctctcgtgggttggaa 1016
DB 388 TONCGGGCAAGGGCGAAATT-ACCTCTCTCCACATGAGAACCTTCTGGGCTTTGGAA 330
QY 1017 acgtgatttgttataaagaatgctcccaagcttggaaagcagggtgggtacatactg 1076
DB 329 ACTGTGATTTGTGTATTAATAAAGTGGCTCCACGCTTGGAAGACGAGGTGGTACAT 270
QY 1077 gagcagcgaaggcgtgagatataaaggaggaggtatgacagcagcagcagcatttc 1136
DB 269 GAGACAGCCCAAGAGCTGATATTAAGAGAGAGGAGATGTGCAAGAACAGAGCGCTTTC 210
QY 1137 ctgggttcgctcccgcttcctcaattctcccttccatcccaaccctcagaacttgat 1196
|||||

Db 209 CTGGTTGGCTCCCGTTCCCTCACTTTCCCTTTTCATTCACCCTAGACTTTGAT 150
Qy 1197 tttaaggatactctgtctctgttccccaagagctccgaattcttgcgtgltgtagatg 1256
Db 149 TTACGGATATCTTCTCTCTGTTCCCATGAGCTCCGATTCCTTGCSTGTGTGTAGATG 90
Qy 1257 aggg 1260
Db 89 AGCG 86

Search completed: May 13, 1999, 19:44:22
Job time: 11327 sec

This Page Blank (uspio)